8-18-2016

Generalizing the GSVD

Mark Gockenbach
Michigan Technological University

Follow this and additional works at: https://digitalcommons.mtu.edu/math-fp

Part of the Mathematics Commons

SIAM J. NUMER. ANAL.
Vol. 54, No. 4, pp. 2517–2540

GENERALIZING THE GSVD*

MARK S. GOCKENBACH†

Abstract. The generalized singular value decomposition (GSVD) of a pair of matrices is the natural tool for certain problems defined on Euclidean space, such as certain weighted least-squares problems, the result of applying Tikhonov regularization to such problems (sometimes called regularization with seminorms), and equality-constrained least-squares problems. There is an extension of the GSVD to pairs of bounded linear operators defined on Hilbert space that turns out to be a natural representation for analyzing the same problems in the infinite-dimensional setting.

Key words. singular value decomposition, Hilbert space, weighted least-squares, Tikhonov regularization

AMS subject classifications. 65J22, 47A52

DOI. 10.1137/15M1019453

1. Introduction.

One of the most useful constructs in numerical analysis is the singular value decomposition (SVD) of a matrix (see, for instance, [5] or [12]). Given any matrix $A \in \mathbb{R}^{m\times n}$, there exist orthogonal matrices $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n\times n}$ and a diagonal matrix $\Sigma \in \mathbb{R}^{m\times n}$ such that

$$A = U\Sigma V^T$$

and the diagonal entries $\sigma_1, \sigma_2, \ldots, \sigma_{\min\{m,n\}}$ satisfy $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{\min\{m,n\}} \geq 0$. If we regard A as defining a linear operator mapping $\mathbb{R}^n$ into $\mathbb{R}^m$, then the columns of V and U define orthonormal bases for the domain and co-domain, respectively, of this operator. In the new variables defined by these bases, the operator is represented by the diagonal matrix Σ, which renders many questions transparent. In particular, the least-squares problem (given $b \in \mathbb{R}^m$, find $x \in \mathbb{R}^n$ to minimize the residual $\|Ax-b\|_2$, where $\|\cdot\|_2$ represents the Euclidean norm) is easily solved using the SVD of A. When A fails to have full rank, so that the solution to the least-squares problem is not unique, one approach to selecting one of the infinitely many least-squares solutions is to choose the one with the smallest Euclidean norm. This *minimum-norm least-squares* solution is also easily identified using the SVD of A.

Certain least-squares problems are defined by two matrices rather than one. For instance, instead of the minimum-norm least-squares solution of $Ax = b$, it is often preferable to compute the solution of

$$\begin{aligned} &\min \|Bx\|_2^2 \\ &\text{ s.t. } x \text{ is a least-squares solution of } Ax = b, \end{aligned} \tag{1.1}$$

where $B \in \mathbb{R}^{p\times n}$ is another matrix. In cases when the solution of (1.1) is not well-determined from a numerical point of view, it is common to apply *Tikhonov regularization* and replace (1.1) by the problem

$$\min_{x\in\mathbb{R}^n} \|Ax-b\|_2^2 + \lambda\|Bx\|_2^2, \tag{1.2}$$

*Received by the editors May 1, 2015; accepted for publication (in revised form) June 21, 2016; published electronically August 18, 2016.
http://www.siam.org/journals/sinum/54-4/M101945.html
†Department of Mathematical Sciences, Michigan Technological University, Houghton, MI 49931-1295 (msgocken@mtu.edu).

where λ a small positive number known as the regularization parameter. As one more example, we mention the equality-constrained least-squares problem,

$$\begin{aligned} &\min \|Ax - b\|_2^2 \\ &\text{s.t. } Bx = d, \end{aligned} \tag{1.3}$$

where now $b \in \mathbb{R}^m$, $d \in \mathbb{R}^p$ are given.

To analyze problems such as (1.1), (1.2), and (1.3), the generalized singular value decomposition (GSVD) is useful. We now present one version of the GSVD [1, Theorem 22.2] (see also [7], [13]). The notation $\mathcal{N}(A)$ represents the null space of the matrix A.

THEOREM 1.1. *Let $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{p\times n}$ be given, and suppose $m \geq n$ and $\mathcal{N}(A) \cap \mathcal{N}(B) = \{0\}$. Then there exist a nonsingular matrix $W \in \mathbb{R}^{n\times n}$, matrices $U \in \mathbb{R}^{m\times n}$, $V \in \mathbb{R}^{p\times p}$ with orthonormal columns, and diagonal matrices $S \in \mathbb{R}^{n\times n}$, $M \in \mathbb{R}^{p\times n}$ such that*

$$A = USW^{-1},\ B = VMW^{-1}.$$

Moreover, the diagonal entries of S and M are nonnegative and

$$S^TS + M^TM = I.$$

The condition that $S^TS+M^TM = I$ means that the diagonal entries $s_1, s_2, \ldots, s_n$ of S and $m_1, m_2, \ldots, m_p$ of M satisfy

$$\begin{aligned} &s_i^2 + m_i^2 = 1,\ i = 1, 2, \ldots, p, \\ &s_i = 1,\ i = p+1, p+2, \ldots, n \end{aligned}$$

(assuming for convenience that $n \geq p$).

For the purposes of analysis (determining existence and uniqueness of solutions, analyzing the conditioning of problems, analyzing convergence of algorithms, and so forth), the GSVD is often the tool of choice for problems such as (1.1), (1.2), and (1.3).

The purpose of this paper is to derive an analogous representation for general linear operators defined on Hilbert spaces and indicate how this representation can be used to analyze the infinite-dimensional analogues of (1.1), (1.2), and (1.3). In the representation to be presented, multiplication by a diagonal matrix on Euclidean space is generalized to multiplication by a bounded measurable function on $L^2(\mu)$, where $(M, \mu, \mathcal{A})$ is a measure space and $L^2(\mu)$ denotes the space of square-integrable functions defined on M. To be specific, if $\theta : M \to \overline{\mathbb{R}}$ (where $\overline{\mathbb{R}}$ denotes the set of extended real numbers) is an essentially bounded measurable function, then we define the multiplication operator $m_\theta : L^2(\mu) \to L^2(\mu)$ by

$$(m_\theta f)(t) = \theta(t)f(t) \text{ for all } t \in M.$$

The SVD can be extended to a singular value expansion (SVE) of a general bounded linear operator defined on Hilbert spaces. We will provide a proof based on the well-known polar decomposition of an operator. In the following theorem, we use the concept of a *partial isometry*. If X and Y are Hilbert spaces and $U : X \to Y$ is a bounded linear operator, then U is called a partial isometry if it defines an isometry on the orthogonal complement of its null space.

THEOREM 1.2. *Let X and Y be real Hilbert spaces and let $T : X \to Y$ be a bounded linear operator. Then there exist a measure space $(M, \mathcal{A}, \mu)$, an isometry $V : L^2(\mu) \to \mathcal{N}(T)^\perp$, an essentially nonnegative, bounded, measurable function $\sigma : M \to [0, \infty)$, and a partial isometry $U : L^2(\mu) \to \overline{\mathcal{R}(T)}$ such that*

$$T = U m_\sigma V^{-1}. \tag{1.4}$$

Moreover, $\mathcal{N}(U) = \mathcal{N}(m_\sigma)$ (and hence $U|_{\mathcal{N}(m_\sigma)^\perp}$ is an isometry).

If $\mathcal{N}(T) = \{0\}$, then U itself is an isometry.

Proof. The polar decomposition theorem [10] shows that T can be written as $T = WA$, where $W : X \to Y$ is a partial isometry and $A : X \to X$ is a positive self-adjoint operator. Moreover, $\mathcal{N}(W) = \mathcal{R}(A)^\perp = \mathcal{N}(A^*) = \mathcal{N}(A) = \mathcal{N}(T)$. By the spectral theorem [6], there exist a measure space $(M, \mathcal{A}, \mu)$, an isometry $V : L^2(\mu) \to X$, and a bounded measurable function $\sigma : M \to \mathbb{R}$ such that

$$A = V m_\sigma V^{-1}.$$

Since A is positive, it is easy to see that σ must be nonnegative. We define $U : L^2(\mu) \to Y$ by $U = WV$. For $R = \mathcal{R}(A)$, we have $\|Wx\|_Y = \|x\|_X$ for all $x \in R$ and $Wx = 0$ for all $x \in R^\perp$. Therefore, if $S = V^{-1}(R)$, then

$$\|Uf\|_Y = \|WVf\|_Y = \|Vf\|_X = \|f\|_{L^2(\mu)} \text{ for all } f \in S$$

and

$$Uf = WVf = 0 \text{ for all } f \in S^\perp \text{ (since } V(S^\perp) = R^\perp \text{ because } V \text{ is unitary).}$$

This shows that $T = U m_\sigma V^{-1}$, U is a partial isometry, and $\mathcal{N}(U) = \mathcal{N}(m_\sigma)$. ◻

In the special case that $T : X \to Y$ is compact, the SVE can be written as

$$T = \sum_{k=1}^{\infty} \sigma_k \psi_k \otimes \phi_k, \tag{1.5}$$

where $\{\phi_k\}$ is an orthonormal sequence in X, $\{\psi_k\}$ is an orthonormal sequence in Y, and $\sigma_1 \geq \sigma_2 \geq \cdots$ are positive numbers converging to zero. Moreover, $\{\phi_k\}$ is a complete orthonormal set for $\mathcal{N}(T)^\perp$ and $\{\psi_k\}$ is a complete orthonormal set for $\overline{\mathcal{R}(T)}$.

We will now show that (1.5) is an example of Theorem 1.2. Let us define $M = \mathbb{Z}^+$, $\mathcal{A} = \mathcal{P}(\mathbb{Z}^+)$ (the power set of $\mathbb{Z}^+$), and μ to be counting measure (that is, for an $E \subset \mathbb{Z}^+$, $\mu(E)$ is the cardinality of E). Then $L^2(\mu)$ is the space of square summable sequences of real numbers (usually denoted by ℓ^2) and, for $\alpha = \{\alpha_k\} \in L^2(\mu)$,

$$\int \alpha^2 = \sum_{k=1}^{\infty} \alpha_k^2.$$

We define $V : L^2(\mu) \to \mathcal{N}(T)^\perp$ by

$$V(\alpha_k) = \sum_{k=1}^{\infty} \alpha_k \phi_k.$$

Then it is straightforward to verify that V is an isometry and that $V^{-1} = V^*$ is defined by

$$\left(V^{-1}(x)\right)_k = \langle \phi_k, x \rangle_X, \ k = 1, 2, \ldots.$$

The sequence $\sigma = \{\sigma_k\}$ is bounded and measurable with respect to the measure space $(M, \mathcal{A}, \mu)$ and

$$m_\sigma \alpha = \{\sigma_k \alpha_k\}.$$

Finally, $U : L^2(\mu) \to Y$ is defined by

$$U\alpha = \sum_{k=1}^{\infty} \alpha_k \psi_k.$$

Therefore, for each $x \in X$, we have

$$\begin{aligned} Um_\sigma V^{-1}x = \sum_{k=1}^{\infty} \left(m_\sigma V^{-1}x\right)_k \psi_k &= \sum_{k=1}^{\infty} \sigma_k \left\langle \phi_k, x \right\rangle_X \psi_k \\ &= \left(\sum_{k=1}^{\infty} \sigma_k \psi_k \otimes \phi_k\right) x \\ &= Tx. \end{aligned}$$

This shows that $T = Um_\sigma V^{-1}$ and also that $Um_\sigma V^{-1}$ is just another way of writing (1.5), the usual SVE of T.

In the next section, we derive an analogue of the GSVD for pairs of linear operators. Section 3 describes three applications in which this expansion is useful. In an important special case, we can produce the explicit form of the expansion and approximate it numerically using a Galerkin algorithm. This is presented in section 4.

2. A generalized SVE for general linear operators. To derive our main result, we will need the following lemma.

LEMMA 2.1. *Let $(M, \mathcal{A}, \mu)$ be a measure space and let $\theta : M \to [0, \infty]$ be a measurable function that is positive and finite a.e. Define*

$$S = \left\{ f \in L^2(\mu) \,:\, \theta^{-1} f \in L^2(\mu) \right\}.$$

Then S is dense in $L^2(\mu)$.

Proof. Let $f \in L^2(\mu)$ be given. For any $\epsilon > 0$, since f^2 is integrable, there exists a measurable subset N_ϵ of M such that $\mu(N_\epsilon) < \infty$ and

$$\int_{M \setminus N_\epsilon} f^2 < \epsilon.$$

(To see this, note that Theorems 2.10 and 2.14 of [3] imply that there is a simple function $g : M \to [0, \infty)$ such that $\left|\int g - \int f^2\right| < \epsilon$. Define N_ϵ to be the support of g. It is clear that $\mu(N_\epsilon) < \infty$, since otherwise g could not be integrable.) Now let $\epsilon > 0$ be given. Define, for each $k \in \mathbb{Z}^+$, $E_k = \{x \in M : \theta(x) < 1/k\} \cap N_{\epsilon/2}$, $F_k = (M \setminus E_k) \cap N_{\epsilon/2}$, and $f_k : M \to [0, \infty)$ by $f_k = f\chi_{F_k}$ (where χ_A is the indicator function of $A \in \mathcal{A}$). We wish to show that $f_k \in S$ for all S and, for k sufficiently large, $\|f_k - f\|_{L^2(\mu)} < \epsilon$.

It is obvious that $f_k \in L^2(\mu)$. We have

$$\int \left(\theta^{-1} f_k\right)^2 = \int \theta^{-2} f^2 \chi_{F_k} = \int_{F_k} \theta^{-2} f^2 \le k^2 \int_{F_k} f^2 \le k^2 \int f^2 < \infty$$

(since $\theta \ge 1/k$ on F_k). This shows that $\theta^{-1} f_k \in L^2(\mu)$, that is, $f_k \in S$.

Finally, $M = F_k \cup E_k \cup (M \setminus N_{\epsilon/2})$ and therefore

$$\begin{aligned}\int (f_k - f)^2 &= \int_{F_k} (f_k - f)^2 + \int_{E_k} (f_k - f)^2 + \int_{M\setminus N_{\epsilon/2}} (f_k - f)^2 \\ &= \int_{F_k} (f - f)^2 + \int_{E_k} (0 - f)^2 + \int_{M\setminus N_{\epsilon/2}} (0 - f)^2 \\ &= \int_{E_k} f^2 + \int_{M\setminus N_{\epsilon/2}} f^2.\end{aligned}$$

The second integral is less than $\epsilon/2$ by construction of $N_{\epsilon/2}$. Moreover, $A \mapsto \int_A f^2$ defines a measure on $\mathcal{A}$ (see Proposition 2.13 of [3]). Since $E_{k+1} \subset E_k$ for all k and each E_k has finite measure, we know that

$$\int_{E_k} f^2 \to \int_E f^2 = 0,$$

where $E = \cap_{n=1}^{\infty} E_k = \{x \in M \,:\, \theta(x) = 0\}$ and $\mu(E) = 0$ by assumption. This implies that

$$\int_{E_k} f^2 < \frac{\epsilon}{2}$$

for all k sufficiently large and hence that

$$\int (f_k - f)^2 < \epsilon$$

for all k sufficiently large. Thus $f_k \to f$ in $L^2(\mu)$ and we have shown that S is dense in $L^2(\mu)$. ◻

We can now state and prove the main theorem of this paper.

THEOREM 2.2. *Let X, Y, and Z be real Hilbert spaces and assume that $T : X \to Y$, $L : X \to Z$ are bounded linear operators. Assume also that there exists $\gamma > 0$ such that*

$$\langle (T^*T + L^*L)x, x\rangle_X \geq \gamma \|x\|_X^2 \text{ for all } x \in X. \tag{2.1}$$

Then the following hold:

- *There exist measure spaces $(M_a, \mathcal{A}_a, \mu_a)$, $(M_b, \mathcal{A}_b, \mu_b)$, and $(M_0, \mathcal{A}_0, \mu_0)$ such that X is isomorphic to $L^2(\mu)$, where $(M, \mathcal{A}, \mu)$ is the direct sum of $(M_a, \mathcal{A}_a, \mu_a)$, $(M_b, \mathcal{A}_b, \mu_b)$, and $(M_0, \mathcal{A}_0, \mu_0)$.*
- *There exists an isomorphism $W : L^2(\mu) \to X$ under which $L^2(\mu_a)$ is isomorphic to $\mathcal{N}(T)$, $L^2(\mu_b)$ is isomorphic to $\mathcal{N}(L)$, and $L^2(\mu_0)$ is isomorphic to*
  $$X_0 = \left\{x \in X \,:\, L^*Lx \in \mathcal{N}(T)^{\perp} \text{ and } T^*Tx \in \mathcal{N}(L)^{\perp}\right\}.$$
- *There exist a partial isometry $U : L^2(\mu) \to Y$ and a bounded measurable function $a : M \to [0, \infty)$ such that*
  $$T = U m_a W^{-1}.$$
  Moreover, $a = 0$ on M_a, $\mathcal{N}(U) = L^2(\mu_a)$, and U defines an isometry from $L^2(\mu_b) \oplus L^2(\mu_0)$ onto $\overline{\mathcal{R}(T)}$.

- *There exist a partial isometry $V : L^2(\mu) \to Z$ and a bounded measurable function $b : M \to [0,\infty)$ such that*

$$L = V m_b W^{-1}.$$

Moreover, $b = 0$ on M_b, $\mathcal{N}(V) = L^2(\mu_b)$, and V defines an isometry from $L^2(\mu_a) \oplus L^2(\mu_0)$ onto $\overline{\mathcal{R}(L)}$.

- *Finally, a and b satisfy $a^2 + b^2 = 1$ on M.*

Proof. The spectral theorem for self-adjoint operators implies that there exist a measure space $(M_1, \mathcal{A}_1, \mu_1)$, an isometry $U_1 : L^2(\mu_1) \to X$, and a bounded measurable function $\theta : M_1 \to \mathbb{R}$ such that

$$T^*T + L^*L = U_1 m_\theta U_1^{-1}.$$

Moreover, since $\langle (T^*T + L^*L)x, x\rangle_X \geq 0$ for all $x \in X$, we know that $\theta \geq 0$. We will show that there exists $\delta > 0$ such that $\theta \geq \delta$ a.e. in M_1. If this is not true, then, for all $k \in \mathbb{Z}^+$, $\mu_1(E_k) > 0$, where

$$E_k = \{x \in M_1 \,:\, \theta(x) \leq 1/k\}.$$

But then

$$\begin{aligned}
\langle (T^*T + L^*L)\, U_1 \chi_{E_k}, U_1 \chi_{E_k}\rangle_X &= \left\langle U_1^{-1}\, (T^*T + L^*L)\, U_1 \chi_{E_k}, \chi_{E_k}\right\rangle_{L^2(\mu_1)} \\
&= \langle m_\theta \chi_{E_k}, \chi_{E_k}\rangle_{L^2(\mu_1)} \\
&\leq \frac{1}{k}\, \langle \chi_{E_k}, \chi_{E_k}\rangle_{L^2(\mu_1)} \\
&= \frac{1}{k}\, \langle U_1 \chi_{E_k}, U_1 \chi_{E_k}\rangle_X\,,
\end{aligned}$$

where $U_1 \chi_{E_k} \neq 0$ in X. Since this holds for all $k \in \mathbb{Z}^+$, we obtain a contradiction to (2.1). This shows that there exists $\delta > 0$ such that $\theta \geq \delta$ a.e. in M_1. It follows that m_σ, where $\sigma = \sqrt{\theta}$, is an invertible operator with a bounded inverse $m_{\sigma^{-1}}$.

Let us define $W_1 : L^2(\mu_1) \to X$ by $W_1 = U_1 m_{\sigma^{-1}}$. Notice that

$$W_1^*\, (T^*T + L^*L)\, W_1 = m_{\sigma^{-1}} U_1^{-1}\, (T^*T + L^*L)\, U_1 m_{\sigma^{-1}} = m_{\sigma^{-1}} m_\theta m_{\sigma^{-1}} = m_1,$$

where m_1 is the multiplication operator on $L^2(\mu_1)$ defined by the constant function 1 (m_1 is the identity operator on $L^2(\mu_1)$).

Next, we define $A : L^2(\mu_1) \to L^2(\mu_1)$, $B : L^2(\mu_1) \to L^2(\mu_1)$ by $A = W_1^* T^* T W_1$, $B = W_1^* L^* L W_1$, respectively. Since A is a bounded self-adjoint operator on $L^2(\mu_1)$ with

$$\langle Af, f\rangle_{L^2(\mu_1)} \geq 0 \text{ for all } f \in L^2(\mu_1),$$

there exist a measure space $(M, \mathcal{A}, \mu)$, an isometry $Q : L^2(\mu) \to L^2(\mu_1)$, and a bounded measurable function $a : M \to [0,\infty)$ such that $Q^{-1} A Q = m_{a^2}$. We now define an operator $W : L^2(\mu) \to X$ by $W = W_1 Q$; notice that W is an isomorphism because both W_1 and Q are isomorphisms. Moreover,

$$W^* T^* T W = Q^{-1} W_1^* T^* T W_1 Q = Q^{-1} A Q = m_{a^2}$$

and

$$W^*\, (T^*T + L^*L)\, W = Q^{-1} W_1^*\, (T^*T + L^*L)\, W_1 Q = Q^{-1} m_1 Q = m_1$$

(in this last series of equations, m_1 first represents the identity on $L^2(\mu_1)$ and then the identity on $L^2(\mu)$). But then

$$\begin{aligned} W^*L^*LW = W^*\left(T^*T + L^*L - T^*T\right)W &= W^*\left(T^*T + L^*L\right)W - W^*T^*TW \\ &= m_1 - m_{a^2} \\ &= m_{1-a^2}. \end{aligned}$$

Since $\langle W^*L^*LWf, f\rangle_{L^2(\mu)} \geq 0$ for all $f \in L^2(\mu)$, it follows that $1 - a^2 \geq 0$ must hold a.e. and hence we can define a bounded measurable function $b : M \to [0,\infty)$ by $b^2 = 1 - a^2$.

We now define

$$\begin{aligned} &M_a = \{x \in M \,:\, a(x) = 0\},\ \mathcal{A}_a = \{E \cap M_a \,:\, E \in \mathcal{A}\},\ \mu_a(G) = \mu(G) \text{ for all } G \in \mathcal{A}_a, \\ &M_b = \{x \in M \,:\, b(x) = 0\},\ \mathcal{A}_b = \{E \cap M_b \,:\, E \in \mathcal{A}\},\ \mu_b(G) = \mu(G) \text{ for all } G \in \mathcal{A}_b, \\ &M_0 = M \setminus (M_a \cup M_b),\ \mathcal{A}_0 = \{E \cap M_0 \,:\, E \in \mathcal{A}\},\ \mu_0(G) = \mu(G) \text{ for all } G \in \mathcal{A}_0. \end{aligned}$$

Since $a^2 + b^2 = 1$, it follows that M_a, M_b, and M_0 partition M into disjoint sets. Therefore, $(M, \mathcal{A}, \mu)$ is (isomorphic to) the direct of the measure spaces $(M_a, \mathcal{A}_a, \mu_a)$, $(M_b, \mathcal{A}_b, \mu_b)$, and $(M_0, \mathcal{A}_0, \mu_0)$, and $L^2(\mu)$ is (isomorphic to) the direct sum of $L^2(\mu_a)$, $L^2(\mu_b)$, and $L^2(\mu_0)$.

It is now easy to show that W defines a bijection from $L^2(\mu_a)$ onto $\mathcal{N}(T)$ and a bijection from $L^2(\mu_b)$ onto $\mathcal{N}(L)$. For example, if $f \in L^2(\mu_a)$ (regarded as a subspace of $L^2(\mu)$), then we have

$$\|TWf\|_Y^2 = \langle f, W^*T^*TWf\rangle_{L^2(\mu)} = \langle f, m_{a^2} f\rangle_{L^2(\mu)} = 0$$

because $a = 0$ on M_a and $f = 0$ on $M \setminus M_a$. Therefore, $Wf \in \mathcal{N}(T)$ for all $f \in L^2(\mu_a)$. Conversely, if $x \in \mathcal{N}(T)$, then

$$\begin{aligned} 0 = \|Tx\|_Y = \langle x, T^*Tx\rangle_X = \left\langle x, W^{-*} m_{a^2} W^{-1}x\right\rangle_X &= \left\langle W^{-1}x, m_{a^2} W^{-1}x\right\rangle_{L^2(\mu)} \\ &= \int_{M_b} (aW^{-1}x)^2 + \int_{M_0} (aW^{-1}x)^2. \end{aligned}$$

This implies that $aW^{-1}x = 0$ a.e. on $M_b \cup M_0$ and hence, since $a > 0$ on $M_b \cup M_0$, that $W^{-1}x = 0$ on $M_b \cup M_0$. This shows that $W^{-1}x \in L^2(\mu_a)$. It follows that $f \in L^2(\mu_a)$ if and only if $Wf \in \mathcal{N}(T)$. The proof that $f \in L^2(\mu_b)$ if and only if $Wf \in \mathcal{N}(L)$ is similar.

Next, we must show that $f \in L^2(\mu_0)$ if and only if $Wf \in X_0$. First, suppose that $f \in L^2(\mu_0)$. Then, for any $u \in \mathcal{N}(L)$,

$$\langle u, T^*TWf\rangle_X = \langle T^*Tu, Wf\rangle_X = \langle (T^*T + L^*L)u, Wf\rangle_X$$

(because $L^*Lu = 0$). We know that $W^*(T^*T + L^*L)W$ is the identity on $L^2(\mu)$, which implies that $T^*T + L^*L = W^{-*}W^{-1}$. It follows that

$$\langle u, T^*TWf\rangle_X = \left\langle W^{-*}W^{-1}u, Wf\right\rangle_X = \left\langle W^{-1}u, f\right\rangle_{L^2(\mu)} = 0$$

because $W^{-1}u \in L^2(\mu_b)$ and $L^2(\mu_b)$, $L^2(\mu_0)$ are orthogonal subspaces of $L^2(\mu)$. Therefore,

$$u \in \mathcal{N}(L) \ \Rightarrow\ \langle u, T^*TWf\rangle_X = 0,$$

and hence $T^*TWf \in \mathcal{N}(L)^\perp$. The proof that $L^*LWf \in \mathcal{N}(T)^\perp$ is similar, and we see that $Wf \in X_0$ for all $f \in L^2(\mu_0)$. Conversely, suppose $x \in X_0$. Consider any $f \in L^2(\mu_a)$. We know that $Wf \in \mathcal{N}(T)$ and hence

$$\begin{aligned}\langle Wf, L^*Lx\rangle_X = 0 \;\Rightarrow\; \langle L^*LWf, x\rangle_X = 0 &\Rightarrow \langle (T^*T + L^*L)Wf, x\rangle_X = 0\\ &\Rightarrow \left\langle W^{-*}W^{-1}Wf, x\right\rangle_X = 0\\ &\Rightarrow \left\langle f, W^{-1}x\right\rangle_X = 0.\end{aligned}$$

Since f was an arbitrary element of $L^2(\mu_a)$, this shows that $W^{-1}x \in L^2(\mu_a)^\perp$. A similar argument shows that $W^{-1}x \in L^2(\mu_b)^\perp$, and hence

$$W^{-1}x \in \left(L^2(\mu_a)^\perp \oplus L^2(\mu_b)^\perp\right)^\perp = L^2(\mu_0).$$

Thus we have shown that $Wf \in X_0$ if and only if $f \in L^2(\mu_0)$.

Now let us define

$$S_a = L^2(\mu_a) \oplus \left\{f \in L^2(\mu_b) \oplus L^2(\mu_0) \,:\, a^{-1}f \in L^2(\mu)\right\}.$$

(Here it is understood that $a^{-1}f = 0$ on M_a.) An argument similar to the proof of Lemma 2.1 shows that S_a is dense in $L^2(\mu)$. Define $U : S_a \to Y$ by $Uf = TWm_{a^{-1}}f$. Then $Uf = 0$ for $f \in L^2(\mu_a)$ and, for $f \in L^2(\mu_b) \oplus L^2(\mu_0)$,

$$\begin{aligned}\|Uf\|_Y^2 = \langle TWm_{a^{-1}}f, TWm_{a^{-1}}f\rangle_Y &= \langle f, m_{a^{-1}}W^*T^*TWm_{a^{-1}}f\rangle_{L^2(\mu)}\\ &= \langle f, m_{a^{-1}}m_{a^2}m_{a^{-1}}f\rangle_{L^2(\mu)}\\ &= \langle f, f\rangle_{L^2(\mu)}\\ &= \|f\|_{L^2(\mu)}^2.\end{aligned}$$

This shows that U is bounded on the dense subspace S_a and satisfies $\|Uf\|_Y = \|f\|_{L^2(\mu)}$ for all f in $S_a \cap \left(L^2(\mu_b) \oplus L^2(\mu_0)\right)$. It follows that U can be extended to a bounded linear operator (still denoted by U) defined on all of $L^2(\mu)$ and satisfying $\|Uf\|_Y = \|f\|_{L^2(\mu)}$ for all $f \in L^2(\mu_a) \oplus L^2(\mu_0)$.

We now wish to show that

$$Um_aW^{-1}x = Tx \text{ for all } x \in X.$$

We know that W maps $L^2(\mu_a)$ onto $\mathcal{N}(T)$ and $L^2(\mu_b) \oplus L^2(\mu_0)$ onto $\mathcal{N}(L) + X_0$; moreover, X is the (nonorthogonal) direct sum of $\mathcal{N}(T)$ and $(\mathcal{N}(L) + X_0)$. Therefore, it suffices to prove that $Um_aW^{-1}x = Tx$ for all $x \in \mathcal{N}(T)$ and for all $x \in \mathcal{N}(L) + X_0$. For any $x \in \mathcal{N}(T)$, we have $W^{-1}x \in L^2(\mu_a) \subset S_a$ and therefore $m_aW^{-1}x = 0$ in M because $a = 0$ on M_a and $W^{-1}x = 0$ on $M_b \cup M_0$. Thus

$$Um_aW^{-1}x = U0 = 0 = Tx.$$

If $x \in \mathcal{N}(L) + X_0$, it is easy to see that $m_aW^{-1}x \in S_a \cap L^2(\mu_b) \oplus L^2(\mu_0)$ and hence

$$Um_aW^{-1}x = TWm_{a^{-1}}m_aW^{-1}x = TWW^{-1}x = Tx.$$

This completes the proof that $Um_aW^{-1} = T$.

It remains only to show that $\mathcal{R}(U) = \overline{\mathcal{R}(T)}$. By definition, $\mathcal{R}(U) \subset \overline{\mathcal{R}(T)}$. Given any $y \in \overline{\mathcal{R}(T)}$, we can find a sequence $\{f_n\} \subset L^2(\mu_b) \oplus L^2(\mu_0)$ such that $TWf_n \to y$. Then

$$\{TWf_n\} = \{Um_aW^{-1}Wf_n\} = \{Um_af_n\}$$

is a Cauchy sequence and therefore, since U is an isometry on $L^2(\mu_b) \oplus L^2(\mu_0)$, it follows that $\{m_a f_n\}$ is a Cauchy sequence in $L^2(\mu)$. Suppose $m_a f_n \to f$. Then

$$Uf = \lim_{n\to\infty} U m_a f_n = y,$$

which shows that $y \in \mathcal{R}(U)$. Thus $\mathcal{R}(U) = \overline{\mathcal{R}(T)}$, as desired.

A similar argument shows that $V = LWm_{b^{-1}}$ defines a bounded linear operator mapping a dense subspace of $L^2(\mu)$ into Z. Extending V to all of $L^2(\mu)$, it can be shown that $\|Vf\|_Z = \|f\|_{L^2(\mu)}$ for all $f \in L^2(\mu_a) \oplus L^2(\mu_0)$, $L = Vm_bW^{-1}$, and $\mathcal{R}(V) = \overline{\mathcal{R}(L)}$. This completes the proof. ∎

We will refer to the representation $T = Um_aW^{-1}$, $L = Vm_bW^{-1}$ as the generalized singular value expansion (GSVE) of the pair T, L.

3. Applications of the GSVE.

Throughout the following discussion, we will assume that X, Y, and Z are real Hilbert spaces and $T : X \to Y$, $L : X \to Z$ are bounded linear operators. We assume that there exists $\gamma > 0$ such that (2.1) holds and that T, L are represented as $T = Um_aW^{-1}$, $L = Vm_bW^{-1}$, as guaranteed by Theorem 2.2.

In many applications, it is necessary to allow L to be a densely defined unbounded linear operator. In such cases, the usual approach is to define a stronger norm under which the domain $D(L)$ of L is a Hilbert space. But then Theorem 2.2 can be applied to T_L, L, where $T_L : D(L) \to Y$ is the restriction of T to $D(L)$. We will discuss this in more detail in section 4.

3.1. Weighted least-squares.

We first address the problem

$$\begin{aligned} &\min \|Lx\|_Z^2 \\ &\text{s.t. } x \text{ is a least-squares solution of } Tx = y. \end{aligned} \tag{3.1}$$

To analyze this problem, we begin by defining $\overline{y} = \mathrm{proj}_{\overline{\mathcal{R}(T)}} y$, the orthogonal projection of y onto $\overline{\mathcal{R}(T)}$. We have

$$\|Tx - y\|_Y^2 = \|Tx - \overline{y}\|_Y^2 + \|\overline{y} - y\|_Y^2 = \|Um_aW^{-1}x - \overline{y}\|_Y^2 + \|\overline{y} - y\|_Y^2,$$

and since U, restricted to $L^2(\mu_b) \oplus L^2(\mu_0)$ is an isometry, we see that

$$\|Um_aW^{-1}x - \overline{y}\|_Y^2 = \|m_aW^{-1}x - U^*\overline{y}\|_{L^2}^2.$$

Similarly, we have

$$\|Lx\|_Z^2 = \|Vm_bW^{-1}x\|_Z^2 = \|m_bW^{-1}x\|_{L^2}^2.$$

We now apply the change of variables $f = W^{-1}x$ and define $\overline{f} = U^*\overline{y}$ to conclude that (3.1) is equivalent to

$$\begin{aligned} &\min \|m_b f\|_{L^2}^2 \\ &\text{s.t. } f \text{ is a least-squares solution of } m_a f = \overline{f} \end{aligned} \tag{3.2}$$

(that is, x solves (3.1) if and only if $f = W^{-1}x$ solves (3.2)). We note that U^*, restricted to $\mathcal{R}(T)^\perp$, is the zero operator, and therefore $U^*\overline{y} = U^*y$.

We can now analyze (3.2). First, we note that

$$\mathcal{R}(m_a) = \{f \in L^2(\mu_b) \oplus L^2(\mu_0) \,:\, a^{-1}f \in L^2(\mu)\}$$

(here $a^{-1}f = \infty\cdot 0$ on M_a, and we always interpret $\infty\cdot 0$ as 0). By Lemma 2.1, $\mathcal{R}(m_a)$ is dense in $L^2(\mu_b)\oplus L^2(\mu_0)$, and it follows that

$$\inf\{\|m_a f - \overline{f}\|_{L^2}^2 \,:\, f\in L^2(\mu)\} = 0$$

always holds. Therefore, (3.2) is equivalent to

$$\begin{aligned} &\min\ \|m_b f\|_{L^2}^2 \\ &\text{s.t. } m_a f = \overline{f} \end{aligned} \tag{3.3}$$

and $m_a f = \overline{f}$ has a solution if and only if $\overline{f}\in\mathcal{R}(m_a)$.

Second, it is easy to see that m_a, restricted to $L^2(\mu_b)\oplus L^2(\mu_0)$, is injective, and it has a bounded inverse if and only if a is essentially bounded away from zero on M_0 (that is, if and only if there exists $\alpha>0$ such that $a\geq\alpha$ a.e. on M_0; note that $a=1$ a.e. on M_b). If this condition holds, then $\mathcal{R}(m_a) = L^2(\mu_b)\oplus L^2(\mu_0)$, (3.2) has a solution for each $\overline{f}\in L^2(\mu_b)\oplus L^2(\mu_0)$, and $m_{a^{-1}}$ is the bounded inverse of m_a (restricted to $L^2(\mu_b)\oplus L^2(\mu_0)$). If it does not hold, then (3.2) has a solution only for $\overline{f}$ in a dense suspace of $L^2(\mu_b)\oplus L^2(\mu_0)$, and $m_{a^{-1}}$ is densely defined and unbounded.

Third, if $\overline{f}\in\mathcal{R}(m_a)$, then $m_a f=\overline{f}$ has a unique solution in $L^2(\mu_b)\oplus L^2(\mu_0)$, namely, $f = m_{a^{-1}}\overline{f}$. The general solution of $m_a f=\overline{f}$ is $f = m_{a^{-1}}\overline{f}+g$ for any $g\in L^2(\mu_a)$. For any such f, we have

$$m_b f = m_b m_{a^{-1}}\overline{f} + m_b g,$$

which implies, since $m_b m_{a^{-1}}\overline{f}\in L^2(\mu_0)$ and $m_b g\in L^2(\mu_a)$, that

$$\|m_b f\|_{L^2}^2 = \|m_b m_{a^{-1}}\overline{f}\|_{L^2}^2 + \|m_b g\|_{L^2}^2.$$

This is obviously minimized by choosing $g=0$; hence the solution to (3.3) (when one exists) is $f = m_{a^{-1}}\overline{f}$.

In the case that $\mathcal{R}(m_a)$ is a proper subspace of $L^2(\mu_b)\oplus L^2(\mu_0)$, $m_{a^{-1}}$ is a closed operator. To see this, suppose $\{f_n\}\subset\mathcal{R}(M_a)$ satisfies $f_n\to f$, $m_{a^{-1}}f_n\to g$. By construction, $a\leq 1$ a.e. and hence $a^{-1}\geq 1$ a.e. on $M_b\cup M_0$. Therefore

$$\begin{aligned} \|m_{a^{-1}}f_n - g\|_{L^2}^2\to 0 &\Rightarrow \int_{M_b\cup M_0}(a^{-1}f_n-g)^2\to 0\\ &\Rightarrow \int_{M_b\cup M_0} a^{-2}(f_n-ag)^2\to 0\\ &\Rightarrow \int_{M_b\cup M_0}(f_n-ag)^2 \leq \int_{M_b\cup M_0} a^{-2}(f_n-ag)^2\to 0\\ &\Rightarrow \|f_n-ag\|_{L^2}^2\to 0. \end{aligned}$$

Thus $f_n\to ag$, which implies that $f=ag$ (since $f_n\to f$ by assumption). But then $a^{-1}f\in L^2(\mu)$, which shows that $f\in\mathcal{R}(m_a)$ and $g = m_{a^{-1}}f$. This shows that $m_{a^{-1}}$ is closed.

It is now a straightforward matter to interpret these results in terms of the original problem (3.1):

1. If $\mathcal{R}(T)$ is closed, then (3.1) has a unique solution for every $y\in Y$ (because in this case, $\overline{y}$ belongs to $\mathcal{R}(T)$ and therefore $\overline{f}$ belongs to $\mathcal{R}(m_a)$), and this solution is given by $x = Wm_{a^{-1}}U^*y$. Moreover, x depends continuously on y.

2. If $\mathcal{R}(T)$ fails to be closed, then (3.1) has a solution if and only if $y \in \mathcal{R}(T) \oplus \mathcal{R}(T)^\perp$, which is a proper dense subspsace of Y. The solution is still given by $x = Wm_{a^{-1}}U^*y$, but in this case the operator $Wm_{a^{-1}}U^*$ is unbounded and the solution x does not depend continuously on y. However, the solution operator $Wm_{a^{-1}}U^*$ is closed (and densely defined).

3.2. Tikhonov regularization. As we have just seen, the weighted least-squares problem (3.1) can be unstable. One approach to stabilizing the solution process is *Tikhonov regularization*, which involves choosing a small positive number λ and solving

$$\min_{x \in X} \|Tx - y\|_Y^2 + \lambda \|Lx\|_Z^2 \tag{3.4}$$

to obtain an approximate solution of (3.1). We will not attempt a complete analysis but merely show that (3.4) has a unique solution for every $y \in Y$ (even when (3.1) does not, that is, even when $\mathcal{R}(T)$ fails to be closed), that this solution depends continuously on y, and that it converges to the unique solution of (3.1), when that solution exists, as $\lambda \to 0$.

The calculations for (3.4) are similar to those for (3.1). We have

$$\begin{aligned}
\|Tx - y\|_Y^2 + \lambda\|Lx\|_Z^2 &= \|Tx - \overline{y}\|_Y^2 + \|\overline{y} - y\|_Y^2 + \lambda\|Lx\|_Z^2 \\
&= \|Um_aW^{-1}x - \overline{y}\|_Y^2 + \|\overline{y} - y\|_Y^2 + \lambda\|Vm_bW^{-1}x\|_Z^2 \\
&= \|m_aW^{-1}x - U^*\overline{y}\|_{L^2}^2 + \|\overline{y} - y\|_Y^2 + \lambda\|m_bW^{-1}x\|_{L^2}^2 \\
&= \|m_af - \overline{f}\|_{L^2}^2 + \|\overline{y} - y\|_Y^2 + \lambda\|m_bf\|_{L^2}^2,
\end{aligned}$$

where, as before, $\overline{y} = \mathrm{proj}_{\overline{\mathcal{R}(T)}} y$, $\overline{f} = U^*\overline{y} = U^*y$, and we have applied the changes of variables $f = W^{-1}x$. Ignoring the additive constant $\|\overline{y} - y\|_Y^2$, we see that (3.4) is equivalent to

$$\min_{f \in L^2(\mu)} \|m_af - \overline{f}\|_{L^2}^2 + \lambda\|m_bf\|_{L^2}^2. \tag{3.5}$$

The objective function in (3.5) is a convex quadratic in f and the optimality condition is the linear equation

$$(m_a^*m_a + \lambda m_b^*m_b)f = m_a^*\overline{f}.$$

The multiplication operators m_a and m_b are self-adjoint and $m_a^*m_a + \lambda m_b^*m_b$ is the multiplication operator $m_{a^2+\lambda b^2}$. Since $a^2 + \lambda b^2 \geq \min\{1, \lambda\}(a^2 + b^2) = \min\{1, \lambda\}$, we see that $m_{a^2+\lambda b^2}$ is invertible with a bounded inverse. Therefore, (3.5) has the unique solution for each $\overline{f}$, namely,

$$f_{\lambda,\overline{f}} = \frac{a}{a^2 + \lambda b^2}\overline{f} = m\left[\frac{a}{a^2 + \lambda b^2}\right]\overline{f}$$

(where we write $m[\theta]$ for the multiplication operator m_θ when convenient). Moreover, it is clear that $f_{\lambda,\overline{f}}$ depends continuously on $\overline{f}$.

If $\overline{f} \in \mathcal{R}(m_a)$, say, $\overline{f} = m_a f_0$, $f_0 \in L^2(\mu_b) \oplus L^2(\mu_0)$, then $f_0 = m_{a^{-1}}\overline{f}$ is the solution of (3.3) and we have

$$\begin{aligned} f_0 - f_{\lambda,\overline{f}} = m_{a^{-1}}\overline{f} - m\left[\frac{a}{a^2+\lambda b^2}\right]\overline{f} &= m\left[\frac{\lambda b^2}{a(a^2+\lambda b^2)}\right]\overline{f} \\ &= m\left[\frac{\lambda b^2}{a(a^2+\lambda b^2)}\right] a f_0 \\ &= m\left[\frac{\lambda b^2}{a^2+\lambda b^2}\right] f_0. \end{aligned}$$

Notice that $b = 0$ on M_b, while $f_0 \in L^2(\mu_b) \oplus L^2(\mu_0)$, that is, $f_0 = 0$ on M_a. It follows that

$$\|f_0 - f_{\lambda,\overline{f}}\|_{L^2}^2 = \int_{M_0} \frac{\lambda^2 b^4}{(a^2+\lambda b^2)^2} f_0^2.$$

Notice that

$$\frac{\lambda^2 b^4}{(a^2+\lambda b^2)^2} f_0^2 \le f_0^2 \text{ on } M_0$$

and

$$\frac{\lambda^2 b^4}{(a^2+\lambda b^2)^2} f_0^2 \to 0 \text{ as } \lambda \to 0$$

(where the convergence is pointwise). It follows from the dominated convergence theorem that

$$\|f_0 - f_{\lambda,\overline{f}}\|_{L^2}^2 \to 0 \text{ as } \lambda \to 0,$$

that is, $f_{\lambda,\overline{f}} \to f_0$ as $\lambda \to 0$.

As in the case of the weighted least-squares problem, it is straightforward to extend the results obtained above to the original problem:

1. For each $y \in Y$, there exists a unique solution $x_{\lambda,y}$ of (3.4), and $x_{\lambda,y}$ depends continuously on y.
2. If $y \in \mathcal{R}(T) \oplus \mathcal{R}(T)^\perp$, then $x_{\lambda,y} \to x_{0,y}$ as $\lambda \to 0$, where $x_{0,y}$ is the unique solution of (3.1).

3.3. Equality-constrained least-squares. We now consider the equality-constrained least-squares problem

$$\begin{aligned} &\min \|Tx - y\|_Y^2 \\ &\text{s.t. } Lx = z. \end{aligned} \tag{3.6}$$

As before, we have

$$\|Tx - y\|_Y^2 = \|m_a f - \overline{f}\|_{L^2}^2 + \|\overline{y} - y\|_Y^2,$$

where $\overline{y} = \mathrm{proj}_{\overline{\mathcal{R}(T)}} y$, $\overline{f} = U^*\overline{y} = U^*y$, and $f = W^{-1}x$. Clearly z must lie in $\mathcal{R}(L)$ or else (3.6) has no solution. Let us suppose that $z = Vh$, where $h = V^*z \in L^2(\mu_a) \oplus L^2(\mu_0)$. The fact that z lies in $\mathcal{R}(L)$ implies that h must equal $m_b W^{-1}x$ for some x, which implies that $b^{-1}h$ belongs to $L^2(\mu)$ (more specifically, to $L^2(\mu_a) \oplus L^2(\mu_0)$). We have

$$Lx = z \;\Leftrightarrow\; Vm_bW^{-1}x = Vh \;\Leftrightarrow\; m_b f = h,$$

and thus (3.6) is equivalent to

$$\begin{aligned} &\min \|m_a f - \overline{f}\|_{L^2} ^2 \\ &\text{s.t. } m_b f = h. \end{aligned} \tag{3.7}$$

The general solution of the equation $m_b f = h$ is $f = b^{-1}h + g$, where $g \in L^2(\mu_b)$. For such a function f, we have

$$\|m_a f - \overline{f}\|_{L^2}^2 = \left\| m\left[\frac{a}{b}\right] h + m_a g - \overline{f} \right\|_{L^2}^2.$$

Since $m[a/b]h$ is zero on $M_a \cup M_b$, g is zero on $M_a \cup M_0$ and $a = 1$ on M_b, and $\overline{f} = 0$ on M_a, we have

$$\|m_a f - \overline{f}\|_{L^2}^2 = \|g - \overline{f}\|_{L^2(\mu_b)}^2 + \left\| \frac{a}{b} h - \overline{f} \right\|_{L^2(\mu_0)}^2.$$

This shows that (3.7) has a unique solution, obtained by taking $g = \overline{f}|_{M_b}$; that is, the solution of (3.7) is $f = b^{-1}h + \overline{f}|_{M_b}$. We can express $\overline{f}|_{M_b}$ as $\chi_{M_b}\overline{f}$, where χ_{M_b} is the indicator function of the set M_b.

It is now straightforward to express the above results in terms of the original problem:

1. If $z \notin \mathcal{R}(L)$, then (3.6) has no solution.
2. If $z \in \mathcal{R}(L)$, then (3.6) has a unique solution, namely, $x = W m_{b^{-1}} V^* z + W \chi_{M_b} U^* y$.

3.4. Discussion. The results presented in the previous section, while neither novel nor deep, illustrate the usefulness of the GSVE of Theorem 2.2. It provides a concrete representation that is convenient for analysis, and it reduces many questions to elementary measure theory.

The basic optimality condition for (3.1), which is both necessary and sufficient because the problem is convex, is that $L^*Lx \in \mathcal{N}(T)^\perp$. It is therefore of interest to understand the subspace

$$\{x \in X \,:\, L^*Lx \in \mathcal{N}(T)^\perp\},$$

which turns out to be related to the inner product $\langle \cdot, \cdot \rangle_*$ on X defined by

$$\langle x, u \rangle_* = \langle Tx, Tu \rangle_Y + \langle Lx, Lu \rangle_Z \text{ for all } x \in X.$$

This inner product was introduced by Locker and Prenter [11], and more details will be given in the next section. For now, we note that it is easy to show, using (2.1), that this does define an inner product.

If U is a subset of X, then we will write $U^\pm$ for the orthogonal complement of U with respect to $\langle \cdot, \cdot \rangle_*$.

LEMMA 3.1. $\{x \in X \,:\, L^*Lx \in \mathcal{N}(T)^\perp\} = \mathcal{N}(T)^\pm$.

Proof. For a given $x \in X$, we have

$$\begin{aligned} x \in \mathcal{N}(T)^\pm &\Leftrightarrow \langle x, u \rangle_* = 0 \text{ for all } u \in \mathcal{N}(T) \\ &\Leftrightarrow \langle Tx, Tu \rangle_Y + \langle Lx, Lu \rangle_Z = 0 \text{ for all } u \in \mathcal{N}(T) \\ &\Leftrightarrow \langle Lx, Lu \rangle_Z = 0 \text{ for all } u \in \mathcal{N}(T) \text{ (since } Tu = 0 \text{ for } u \in \mathcal{N}(T)) \\ &\Leftrightarrow \langle L^*Lx, u \rangle_X = 0 \text{ for all } u \in \mathcal{N}(T) \\ &\Leftrightarrow L^*Lx \in \mathcal{N}(T)^\perp. \end{aligned}$$

This completes the proof. ∎

The following representation of $\mathcal{N}(T)^{\pm}$ is also of interest.

LEMMA 3.2. *$\mathcal{N}(T)^{\pm} = \mathcal{N}(L) + X_0$, where X_0 is the space defined in Theorem* 2.2*:*

$$X_0 = \left\{ x \in X \; : \; L^* L x \in \mathcal{N}(T)^{\perp} \text{ and } T^* T x \in \mathcal{N}(L)^{\perp} \right\}.$$

Proof. If $x \in \mathcal{N}(L) + X_0$, then it is clear that $L^* L x \in \mathcal{N}(T)^{\perp}$. Suppose, on the other hand, that $x \in X$ and $L^* L x \in \mathcal{N}(T)^{\perp}$. Let $\overline{x}$ be the projection of x onto $\mathcal{N}(L)$ under the $*$-inner product. (Note that L is obviously bounded under the $*$-norm, and hence $\mathcal{N}(L)$ is closed under this norm.) Then

$$L^* L(x - \overline{x}) = L^* L x - L^* L \overline{x} = L^* L x \in \mathcal{N}(T)^{\perp}$$

and

$$\begin{aligned}
&\langle x - \overline{x}, u \rangle_* = 0 \text{ for all } u \in \mathcal{N}(L) \\
\Rightarrow\ &\langle T(x - \overline{x}), Tu \rangle_Y + \langle L(x - \overline{x}), Lu \rangle_Z = 0 \text{ for all } u \in \mathcal{N}(L) \\
\Rightarrow\ &\langle T(x - \overline{x}), Tu \rangle_Y = 0 \text{ for all } u \in \mathcal{N}(L) \text{ (since } Lu = 0 \text{ for } u \in \mathcal{N}(L)) \\
\Rightarrow\ &\langle T^* T(x - \overline{x}), u \rangle_X = 0 \text{ for all } u \in \mathcal{N}(L) \\
\Rightarrow\ &T^* T(x - \overline{x}) \in \mathcal{N}(L)^{\perp}.
\end{aligned}$$

Thus we have shown that $x - \overline{x} \in X_0$ and hence that

$$x = \overline{x} + (x - \overline{x}) \in \mathcal{N}(L) + X_0,$$

as desired. ∎

According to Theorem 2.2, X is isomorphic to $L^2(\mu) = L^2(\mu_a) \oplus L^2(\mu_b) \oplus L^2(\mu_0)$, and X has the decomposition $X = \mathcal{N}(T) + \mathcal{N}(L) + X_0$. This decomposition is not orthogonal with respect to the X-inner product, but it is easy to check that it is an orthogonal decomposition with respect to the $*$-inner product. Moreover, for any $f, g \in L^2(\mu)$, we have

$$\begin{aligned}
\langle f, g \rangle_{L^2} = \int_M fg \, d\mu &= \int_M (a^2 + b^2) fg \, d\mu \\
&= \int_M (af)(ag) \, d\mu + \int_M (bf)(bg) \, d\mu \\
&= \langle m_a f, m_a g \rangle_{L^2} + \langle m_b f, m_b g \rangle_{L^2} \\
&= \langle U m_a f, U m_a g \rangle_Y + \langle V m_b f, V m_b g \rangle_Z \\
&= \langle U m_a W^{-1} W f, U m_a W^{-1} W g \rangle_Y + \langle V m_b W^{-1} W f, V m_b W^{-1} W g \rangle_Z \\
&= \langle TWf, TWg \rangle_Y + \langle LWf, LWg \rangle_Z \\
&= \langle Wf, Wg \rangle_*.
\end{aligned}$$

Therefore, W defines an isometry between $L^2(\mu)$ and X under the $*$-norm; under this isometry, $L^2(\mu_a)$ is mapped to $\mathcal{N}(T)$, $L^2(\mu_b)$ is mapped to $\mathcal{N}(L)$, and $L^2(\mu_0)$ is mapped to X_0. This explains why the GSVE is so convenient for the applications presented in section 3.

As mentioned earlier, in many applications, the operator L is unbounded and densely defined. We can apply the GSVE of Theorem 2.2 by restricting T to the domain $D(L)$ of L and imposing the inner product $\langle \cdot, \cdot \rangle_*$. It can be shown that $D(L)$

is a Hilbert space under this inner product and all of the analysis presented above is valid, with only obvious changes. (For example, the weighted least-squares problem (3.1) has a solution if and only if $\overline{y} \in \mathcal{R}(T_L)$, where $T_L = T|_{D(L)}$.) In this setting, $W : L^2(\mu) \to D(L)$ is an isomorphism (that is, it is continuous with a continuous inverse). Viewed as an operator mapping $L^2(\mu)$ into X (with the X-norm), W is still continuous, although W^{-1} is now unbounded.

4. Computing the GSVE: An important special case. As demonstrated in section 3, Theorem 2.2 provides a powerful tool for the analysis of certain problems; Theorem 1.2 does the same for ordinary (unweighted) least-squares and for Tikhonov regularization. However, the representations of Theorems 1.2 and 2.2 are abstract in the sense that the form of the measure space(s) is not specified. As noted above, when T is compact, its SVE can be written as

$$T = \sum_{n=1}^{\infty} \sigma_n \psi_n \otimes \phi_n, \tag{4.1}$$

where $\{\phi_n\}$ is a complete orthonormal set for $\mathcal{N}(T)^\perp$, $\{\psi_n\}$ is an orthonormal set in Y, and $\{\sigma_n\}$ is a nonincreasing sequence of positive real numbers that converges to zero. (Here we are assuming that T does not have finite rank; in the alternate case, the sum in (4.1) is finite. We will not discuss this degenerate case.)

We will now show that the GSVE of the pair T, L has a similarly explicit representation when T is compact. We will assume that L is densely defined, because this is the most interesting case, although the reader will notice that we never use the assumption that the domain of L is a *proper* subspace of the domain of T.

We assume that X, Y, and Z are real Hilbert spaces, that $T : X \to Y$ is a bounded linear operator, and that $L : D(L) \to Z$ is linear and closed, where $D(L)$ is a dense subspace of X. We assume that there exists $\gamma > 0$ such that

$$\langle Tx, Tx\rangle_Y + \langle Lx, Lx\rangle_Z \geq \gamma \|x\|_X^2 \text{ for all } x \in D(L). \tag{4.2}$$

We use the inner product $\langle \cdot, \cdot\rangle_*$ defined above on $D(L)$ and write $\|\cdot\|_*$ for the corresponding norm. Following Locker and Prenter [11], we will write $T^\#$ for the adjoint of T regarded as an operator defined on $D(L)$:

$$\langle Tx, y\rangle_Y = \left\langle x, T^\# y\right\rangle_* \text{ for all } x \in D(L), y \in Y.$$

(Above we wrote T_L to denote the restriction of T to $D(L)$, but henceforth we will allow the context to determine the domain of T.) Similarly, we write $L^\#$ for the adjoint of L. We will write T^* and L^* for the adjoints when the X-inner product is applied on the domain.

We collect some useful results in the following lemma; the proofs can be found in [11] or [4].

LEMMA 4.1. *Let X, Y, and Z be real Hilbert spaces and assume that $T : X \to Y$ is a bounded linear operator and $L : D(L) \to Z$ is a densely defined and closed linear operator. Assume also that there exists $\gamma > 0$ such that* (4.2) *holds.*

1. *$D(L)$ is a Hilbert space under the $*$-inner product.*
2. *T and L are both bounded as operators defined on $D(L)$ (under the $*$-norm) and their operator norms are bounded by* 1.
3. *$\mathcal{N}(T) \cap D(L)$ and $\mathcal{N}(L)$ are orthogonal with respect to the $*$-inner product.*
4. *$T^*T + L^*L$ defines an invertible linear operator from $D(L^*L)$ onto X, and $(T^*T + L^*L)^{-1}$ is bounded under the $*$-norm on $D(L^*L)$.*

5. $T^{\#} = (T^*T + L^*L)^{-1}T^*$.
6. $L^{\#}|_{D(L^*)} = (T^*T + L^*L)^{-1}L^*$, *and this operator extends to a bounded linear operator defined on all of* Z.
7. $T^{\#}T + L^{\#}L$ *is the identity operator on* $D(L)$.

Using these results, together with the assumption that T is compact, we can derive an explicit representation of the GSVE of T, L. We will need one more notation. For $x \in D(L)$, $y \in Y$, we write $y \otimes_* x$ for the outer product defined by the $*$-inner product:

$$(y \otimes_* x)u = \langle x, u \rangle_* \, y \text{ for all } u \in D(L).$$

THEOREM 4.2. *Let X, Y, and Z be real Hilbert spaces. Assume that $T : X \to Y$ is a compact linear operator and $L : D(L) \to Z$ is a densely defined linear operator. Assume also that there exists $\gamma > 0$ such that* (4.2) *holds. Then there exists a complete orthonormal set $\{\phi_n : n \in I\}$ for $D(L)$, where I is a countable index set, a partition $M_0 \cup M_a \cup M_b$ of I, orthonormal sets $\{\psi_n : n \in M_0 \cup M_b\}$ of Y, $\{\theta_n : n \in M_0 \cup M_a\}$ of Z, and subsets $\{a_n : n \in I\}$, $\{b_n : n \in I\}$ of $\mathbb{R}$ such that*

$$T = \sum_{n \in M_0 \cup M_b} a_n \psi_n \otimes_* \phi_n,$$

$$L = \sum_{n \in M_0 \cup M_a} b_n \theta_n \otimes_* \phi_n,$$

and $0 \le a_n, b_n \le 1$, $a_n^2 + b_n^2 = 1$ *for all* $n \in I$.

Proof. Since T is compact with respect to the X-norm, it is also compact with respect to the $*$-norm. Therefore, $T^{\#}T$ is compact and there exists an orthonormal set $\{\phi_n : n \in E\}$ and a set $\{\lambda_n \in \mathbb{R} : n \in E\}$ such that

$$T^{\#}T = \sum_{n \in E} \lambda_n \phi_n \otimes_* \phi_n.$$

Here E is countable index set and $\{\phi_n : n \in E\}$ is a complete orthonormal set for the orthogonal complement (with respect to $\langle \cdot, \cdot \rangle_*$) of $\mathcal{N}(T) \cap D(L)$. By Lemma 4.1, $\mathcal{N}(L)$ is orthogonal to $\mathcal{N}(T) \cap D(L)$. Moreover, since $T^{\#}T + L^{\#}L$ is the identity,

$$Lx = 0 \;\Rightarrow\; T^{\#}Tx = x,$$

which means that every element of $\mathcal{N}(L)$ is an eigenvector of $T^{\#}T$ corresponding to the eigenvalue 1. Similarly,

$$Tx = 0 \;\Rightarrow L^{\#}Lx = x.$$

We can now choose a complete orthonormal set $\{\phi_n : n \in F\}$ for $\mathcal{N}(T) \cap D(L)$ so that $\{\phi_n : n \in I\}$, where $I = E \cup F$, is a complete orthonormal set for $D(L)$. By the above comments, we can partition I into $M_0 \cup M_b \cup M_a$ ($M_0 \cup M_b = E$, $M_a = F$) so that

$$L^{\#}L\phi_n = \phi_n \text{ and } T\phi_n = 0 \text{ for all } n \in M_a,$$
$$T^{\#}T\phi_n = \phi_n \text{ and } L\phi_n = 0 \text{ for all } n \in M_b,$$
$$T^{\#}T\phi_n = \lambda_n\phi_n \text{ and } 0 < \lambda_n < 1 \text{ for all } n \in M_0.$$

We define $a_n = \sqrt{\lambda_n}$ for $n \in M_0 \cup M_b$ and $a_n = 0$ for $n \in M_a$. Then, for each $n \in I$, $0 \le a_n \le 1$ and

$$L^{\#}L\phi_n = (T^{\#}T + L^{\#}L)\phi_n - T^{\#}T\phi_n = \phi_n - a_n^2\phi_n = (1 - a_n^2)\phi_n.$$

For each n, we define $b_n = \sqrt{1 - a_n^2}$, so that $L^{\#}L\phi_n = b_n^2\phi_n$.

Finally, we define

$$\psi_n = a_n^{-1}T\phi_n,\ n \in M_0 \cup M_b,$$
$$\theta_n = b_n^{-1}L\phi_n,\ n \in M_0 \cup M_a.$$

It is then straightforward to prove that $\{\psi_n : n \in M_0 \cup M_b\}$ and $\{\theta_n : n \in M_0 \cup M_a\}$ are orthonormal sets in Y and Z, respectively (the proof is exactly the same as in the derivation of the finite-dimensional SVD). Also, for any $x \in D(L)$,

$$x = \sum_{n \in I} \langle \phi_n, x \rangle_* \phi_n$$

and hence

$$\begin{aligned} Tx = \sum_{n \in I} \langle \phi_n, x \rangle_* T\phi_n &= \sum_{n \in M_0 \cup M_b} \langle \phi_n, x \rangle_* T\phi_n \text{ (since } T\phi_n = 0 \text{ for } n \in M_a) \\ &= \sum_{n \in M_0 \cup M_b} \langle \phi_n, x \rangle_* a_n\psi_n \\ &= \sum_{n \in M_0 \cup M_b} a_n(\psi_n \otimes_* \phi_n)x. \end{aligned}$$

Therefore,

$$T = \sum_{n \in M_0 \cup M_b} a_n\psi_n \otimes_* \phi_n.$$

A similar calculation shows that

$$L = \sum_{n \in M_0 \cup M_a} b_n\theta_n \otimes_* \phi_n.$$

∎

4.1. A Galerkin scheme for approximating the GSVE. We now propose a Galerkin method for approximately computing the GSVE of Theorem 4.2. Assume that $X_n = \operatorname{span}\{x_i : i = 1, 2, \ldots, n\}$, $Y_m = \operatorname{span}\{y_i : i = 1, 2, \ldots, m\}$, and $Z_p = \operatorname{span}\{z_i : i = 1, 2, \ldots, p\}$ are finite-dimensional subspaces of $D(L)$, Y, Z, respectively. (Since we will be interested in the convergence as $n, m, p \to \infty$ and the bases are allowed to change as n changes, we should write, for example, $x_i^{(n)}$, $i = 1, 2, \ldots, n$, for the basis elements of X_n. However, we will use the simpler notation.)

The convergence of the method to be proposed will follow from the general theory for symmetric, variationally posed eigenvalue problems presented in Boffi's survey article [2]. We begin by posing the eigenvalue problem for $T^{\#}T$ in variational form. To be consistent with [2], we write this eigenvalue problem as

$$T^{\#}T\phi = \lambda^{-1}\phi. \tag{4.3}$$

Then

$$\begin{aligned} T^{\#}T\phi = \lambda^{-1}\phi &\Rightarrow \left\langle T^{\#}T\phi, \psi\right\rangle_* = \lambda^{-1}\left\langle \phi, \psi\right\rangle_* \\ &\Rightarrow \left\langle \phi, \psi\right\rangle_* = \lambda\left\langle T^{\#}T\phi, \psi\right\rangle_* \\ &\Rightarrow \left\langle \phi, \psi\right\rangle_* = \lambda\left\langle T\phi, T\psi\right\rangle_Y \text{ for all } \psi \in D(L). \end{aligned}$$

We now define $a : D(L) \times D(L) \to \mathbb{R}$ by $a(\phi, \psi) = \langle \phi, \psi\rangle_*$ and $b : X \times X \to \mathbb{R}$ by $b(\phi, \psi) = \langle T\phi, T\psi\rangle_Y$. Then (4.3) is equivalent to the variationally posed eigenvalue problem

$$\phi \in D(L),\ a(\phi, \psi) = \lambda b(\phi, \psi) \text{ for all } \psi \in D(L). \tag{4.4}$$

The bilinear form a is $D(L)$-elliptic, and (4.4) fits into the class of problems described by Boffi in [2] (see also [9]). The Galerkin formulation is

$$\tilde{\phi} \in X_n,\ a(\tilde{\phi}, \tilde{\psi}) = \lambda b(\tilde{\phi}, \tilde{\psi}) \text{ for all } \tilde{\psi} \in X_n. \tag{4.5}$$

A straightforward calculation shows that (4.5) is equivalent to

$$\tilde{\phi} = \sum_{j=1}^{n} \alpha_j x_j \text{ and } G\alpha = \lambda M\alpha, \tag{4.6}$$

where $G, M \in \mathbb{R}^{n\times n}$ are defined by $G_{ij} = \langle x_j, x_i\rangle_*$, $M_{ij} = \langle Tx_j, Tx_i\rangle_Y$, and $\alpha \in \mathbb{R}^n$ is a generalized eigenvector of G and M.

Let $\tilde{\lambda}_k$, $\tilde{\phi}_k$, $k = 1, 2, \ldots, n$, be the solutions of (4.5). To compare the eigenvalues of (4.5) with those of (4.4), we have to order them consistently. Let us assume that

$$\tilde{\lambda}_1 \le \tilde{\lambda}_2 \le \cdots \le \tilde{\lambda}_n$$

and that $M_b = \{1, 2, \ldots, r\}$, $M_0 = \{r + 1, r + 2, \ldots\}$, and λ_k is increasing for $k \in M_0$. Since $\lambda_k = a_k^{-2}$, where a_k is defined by Theorem 4.2, it follows that $\lambda_k = 1$ for $k \in M_b$, $\lambda_k > 1$ for $k \in M_0$, and hence that

$$\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n \le \lambda_{n+1} \le \cdots .$$

Theorem 9.12 of [2] implies that there exists a constant c_k (independent of n) such that

$$\lambda_k \le \tilde{\lambda}_k \le \lambda_k + c_k \sup_{\substack{u \in E_{\lambda_k} \\ \|u\|_* = 1}} \inf_{v\in X_n} \|u - v\|_*^2,$$

where E_{λ_k} is the eigenspace corresponding to the eigenvalue λ_k of (4.5). Assuming that $\cup_{n=1}^{\infty} X_n$ is dense in $D(L)$, this implies that $\tilde{\lambda}_k \to \lambda_k$ as $k \to \infty$.

Theorem 9.13 of [2] implies that there exists a constant C_k (independent of n) such that

$$\inf_{\tilde{\psi}\in\tilde{E}_{\lambda_k}} \|\phi_k - \tilde{\psi}\|_* \le C_k \sup_{\substack{u \in E_{\lambda_k} \\ \|u\|_* = 1}} \inf_{v\in X_n} \|u - v\|_*.$$

Here, $\tilde{E}_{\lambda_k}$ is the space spanned by the approximate eigenvectors associated with the approximate eigenvalues converging to λ_k.

We are interested in approximating a_k, b_k rather than λ_k. Note that $\tilde{\lambda} = \infty$ can be a solution of (4.6), signaling that the matrix M is singular and $\tilde{\phi}_k$ belongs to the null space of M. For $\lambda_k \in \mathbb{R}$, we define

$$a_k = \lambda_k^{-1/2},\ \tilde{a}_k = \tilde{\lambda}_k^{-1/2},$$

while we take $a_k = 0$, $b_k = 1$ when $\lambda_k = \infty$. We will write

$$\tilde{\epsilon}_k = \sup_{\substack{u \in E_{\lambda_k} \\ \|u\|_* = 1}} \inf_{v \in X_n} \|u - v\|_*. \tag{4.7}$$

When λ_k is finite, then it is straightforward to show that there exists a constant c'_k such that

$$\lambda_k \le \tilde{\lambda}_k \le \lambda_k + c_k \tilde{\epsilon}_k^2 \ \Rightarrow\ \tilde{a}_k \le a_k \le \tilde{a}_k + c'_k \tilde{\epsilon}_k^2.$$

Similarly, if we define $\tilde{b}_k = \sqrt{1 - \tilde{a}_k^2}$, then we can show that there exists a constant c''_k such that

$$\tilde{b}_k - c''_k \tilde{\epsilon}_k^2 \le b_k \le \tilde{b}_k.$$

Finally, we must approximate ψ_k, θ_k. Since $\psi_k = a_k^{-1} T \phi_k$ for $k \in M_0 \cup M_b$, we define

$$\tilde{\psi}_k = \tilde{a}_k^{-1} P_{Y_n} T \tilde{\phi}_k,\ k = 1, 2, \ldots, n,\ \tilde{a}_k \ne 0,$$

where P_{Y_n} is the orthogonal projector onto Y_n. If any $\tilde{\lambda}_k = \infty$, then the corresponding $\tilde{\phi}_k$ belongs to the null space of T_n.

Similarly, we define

$$\tilde{\theta}_k = \tilde{b}_k^{-1} P_{Z_n} L \tilde{\phi}_k,\ k = r+1, 2, \ldots, n,\ \tilde{b}_k \ne 1,$$

and $\tilde{\theta}_k = P_{Z_n} L \tilde{\phi}_k$ for each k such that $\tilde{\lambda}_k = \infty$. Since T_n converges to T in norm, it is clear that each $\tilde{\psi}_k$ approximates ψ_k with an error bounded by a multiple of $\tilde{\epsilon}_k$, and similarly for $\tilde{\theta}_k$.

We now show how to compute $P_{Y_n} T \phi$ and $P_{Z_n} T \phi$ for $\phi \in X_n$. We define $H \in \mathbb{R}^{m \times m}$ to be the Gram matrix for the basis $\{y_1, y_2, \ldots, y_m\}$ and $J \in \mathbb{R}^{p \times p}$ to be the Gram matrix for the basis $\{z_1, z_1, \ldots, z_p\}$. For $x = \sum_{j=1}^n \alpha_j x_j$, we have

$$Tx = \sum_{j=1}^n \alpha_j T x_j \ \Rightarrow\ \langle Tx, y_i \rangle_Y = \sum_{j=1}^n \langle T x_j, y_i \rangle_Y \alpha_j = (A\alpha)_i,$$

where $a \in \mathbb{R}^{m \times n}$ is defined by $A_{ij} = \langle T x_j, y_i \rangle_Y$. It follows that

$$P_{Y_n} T \left(\sum_{j=1}^n \alpha_j x_j \right) = \sum_{j=1}^m \beta_j y_j,$$

where $\beta = H^{-1} A \alpha$. Similarly,

$$P_{Z_n} L \left(\sum_{j=1}^n \alpha_j x_j \right) = \sum_{j=1}^m \gamma_j z_j,$$

where $\gamma = J^{-1}B\alpha$ and $B \in \mathbb{R}^{p\times n}$ is defined by $B_{ij} = \langle Lx_j, z_i\rangle_Z$. Therefore, if $\alpha^{(1)}, \alpha^{(2)}, \ldots, \alpha^{(n)}$ are the eigenvectors of (4.6), then

$$\tilde{\phi}_k = \sum_{j=1}^{n} \alpha_j^{(k)} x_j$$

and $\tilde{\psi}$, $\tilde{\theta}$ can be computed as indicated above.

Example 4.3. *To construct an example, we define $X = L^2(0,1)$, $T : X \to X$ by*

$$(Tx)(s) = \int_0^1 k(s,t)x(t)\,dt,\ 0 < s < 1,$$

where

$$k(s,t) = \begin{cases} s(1-t), & s \le t, \\ t(1-s), & s > t. \end{cases}$$

Then T is the solution operator of the two-point boundary value problem

$$\begin{aligned} -u'' &= x \text{ in } (0,1), \\ u(0) &= 0, \\ u(1) &= 0. \end{aligned}$$

The operator $L : D(L) \to X$ is defined by $Lx = x'$, where $D(L) = H^1(0,1)$. In this example, $Y = Z = X$.

Given a uniform mesh $\{[0,h],[h,2h],\ldots,[1-h,1]\}$ on $[0,1]$, where $h = 1/n$, we define $X_n = Y_n = Z_n$ to be the space of continuous piecewise linear functions and write $\{x_0, x_1, \ldots, x_n\}$ for the standard nodal basis. Since $T^{\#}T$ maps X into $D(L^*L)$ and $D(L^*L)$ is contained in $H^2(0,1)$ in this case, it follows that each solution of (4.3) belongs to $H^2(0,1)$. From this and standard approximation theory, it follows that $\tilde{\epsilon}_k$, defined by (4.7), is of order h. Therefore, we can expect, with the scheme described above, to approximate a_k, b_k with an error of order h^2 and ϕ_k, ψ_k, θ_k with an error of order h.

In this example, we do not have exact values of a_k, b_k, ϕ_k, ψ_k, θ_k for comparison. However, if $\tilde{a}_k = a_k + O(h^2)$, then there exists a constant γ_k such that $\tilde{a}_{k,h} \approx a_k + \gamma_k h^2$, where we temporarily write $\tilde{a}_{k,h}$ to indicate the dependence of $\tilde{a}_k$ on h. We can now perform the calculation on two meshes, of sizes h and $h/2$; then a simple calculation shows that

$$\gamma_k \approx \frac{4(\tilde{a}_{k,h} - \tilde{a}_{k,h/2})}{3h^2}.$$

In other words, the expression $4(\tilde{a}_{k,h} - \tilde{a}_{k,h/2})/(3h^2)$ should be approximately constant if $\tilde{a}_{k,h}$ converges with asymptotic error $O(h^2)$. Conversely, it is easy to show that this expression is $O(h^{p-2})$ (and hence not approximately constant) if the convergence in $\tilde{a}_k$ is $O(h^p)$ for $p \ne 2$.

Table 1 shows the computed values of $\tilde{a}_k$, for $k = 1,2,3,4,5$, corresponding to meshes with n equal to $20, 40, 80, 160, 320$, and also the computed estimates for γ_k. The expected rate of convergence is observed.

4.2. An alternate computational scheme. The Galerkin method proposed above has at least two drawbacks. First, it is necessary to compute the matrix M defined by $M_{ij} = \langle Tx_j, Tx_i\rangle_Y$. This poses an extra burden over computing the

TABLE 1

Example 4.3: estimates of $\tilde{a}_k$ on meshes with different values of n. Estimates of γ_k are given in parentheses (not shown for $k = 1$ because $\tilde{a}_1 = 1$ to machine precision).

	$n = 20$	$n = 40$	$n = 80$	$n = 160$	$n = 320$
$k = 1$	1.0000	1.0000	1.0000	1.0000	1.0000
$k = 2$	$7.0001 \cdot 10^{-3}$	$7.0072 \cdot 10^{-3}$	$7.0090 \cdot 10^{-3}$	$7.0094 \cdot 10^{-3}$	$7.0095 \cdot 10^{-3}$
		$(-3.8 \cdot 10^{-3})$	$(-3.8 \cdot 10^{-3})$	$(-3.8 \cdot 10^{-3})$	$(-3.8 \cdot 10^{-3})$
$k = 3$	$1.6915 \cdot 10^{-3}$	$1.6979 \cdot 10^{-3}$	$1.6995 \cdot 10^{-3}$	$1.6999 \cdot 10^{-3}$	$1.7000 \cdot 10^{-3}$
		$(-3.4 \cdot 10^{-3})$	$(-3.4 \cdot 10^{-3})$	$(-3.4 \cdot 10^{-3})$	$(-3.4 \cdot 10^{-3})$
$k = 4$	$6.4727 \cdot 10^{-4}$	$6.5248 \cdot 10^{-4}$	$6.5379 \cdot 10^{-4}$	$6.5412 \cdot 10^{-4}$	$6.5420 \cdot 10^{-4}$
		$(-2.8 \cdot 10^{-3})$	$(-2.8 \cdot 10^{-3})$	$(-2.8 \cdot 10^{-3})$	$(-2.8 \cdot 10^{-3})$
$k = 5$	$3.1155 \cdot 10^{-4}$	$3.1589 \cdot 10^{-4}$	$3.1698 \cdot 10^{-4}$	$3.1725 \cdot 10^{-4}$	$3.1732 \cdot 10^{-4}$
		$(-2.3 \cdot 10^{-3})$	$(-2.3 \cdot 10^{-3})$	$(-2.3 \cdot 10^{-3})$	$(-2.3 \cdot 10^{-3})$

matrix A, which is required in any case. In particular, if T is a first-kind integral operator, then A_{ij} is defined by a double integral, while M_{ij} requires the computation of a triple integral. A second drawback is that $\{\tilde{\psi}_1, \tilde{\psi}_2, \ldots, \tilde{\psi}_n\}$ and $\{\tilde{\theta}_1, \tilde{\theta}_2, \ldots, \tilde{\theta}_n\}$ are only approximately orthonormal. (On the other hand, $\{\tilde{\phi}_1, \tilde{\phi}_2, \ldots, \tilde{\phi}_n\}$ is orthonormal in $D(L)$ under the $*$-inner product.)

We now propose a different scheme that avoids the above difficulties. We define $\hat{T} : X_n \to Y_m$ by $\hat{T} = P_{Y_m} T|_{X_n}$ and $\hat{L} : X_n \to Z_p$ by $\hat{L} = P_{Z_p} L|_{X_n}$. We can compute the GSVE of $\hat{T}$, $\hat{L}$ by using the matrix GSVD. We define a discrete version of the $*$-inner product by

$$\langle u, v \rangle_{*_n} = \left\langle \hat{T}u, \hat{T}v \right\rangle_Y + \left\langle \hat{L}u, \hat{L}v \right\rangle_Z \text{ for all } u, v \in X_n.$$

We also write $y \otimes_{*_n} x$ for the outer product defined by the $*_n$-inner product. If $u = \sum_{i=1}^n \alpha_i x_i$ and $v = \sum_{i=1}^n \beta_i x_i$, then it follows from the above calculations that

$$\begin{aligned}
\langle u, v \rangle_{*_n} &= \left\langle \hat{T}\left(\sum_{j=1}^n \alpha_j x_j\right), \hat{T}\left(\sum_{i=1}^n \beta_i x_i\right) \right\rangle_Y + \left\langle \hat{L}\left(\sum_{j=1}^n \alpha_j x_j\right), \hat{L}\left(\sum_{i=1}^n \beta_i x_i\right) \right\rangle_Z \\
&= \left(H^{-1}A\alpha\right) \cdot H\left(H^{-1}A\beta\right) + \left(J^{-1}B\alpha\right) \cdot J\left(J^{-1}B\beta\right) \\
&= \alpha \cdot \left(A^T H^{-1} A + B^T J^{-1} B\right)\beta.
\end{aligned}$$

THEOREM 4.4. *Let $\hat{T}$, $\hat{L}$, A, and B be defined as in the previous paragraph. Assume that*

$$H^{-1/2}A = USW^{-1}, \ J^{-1/2}B = VMW^{-1}$$

is the GSVD of the matrix pair $H^{-1/2}A$, $J^{-1/2}B$ (as defined in Theorem 1.1), and define $\tilde{U} = H^{-1/2}U$, $\tilde{V} = J^{-1/2}V$. Then the GSVE of $\hat{T}$, $\hat{L}$ is defined by

$$\hat{T} = \sum_{i=1}^{\min\{m,n\}} \tilde{a}_i \tilde{\psi}_i \otimes_{*_n} \tilde{\phi}_i, \ \hat{L} = \sum_{i=1}^{\min\{p,n\}} \tilde{b}_i \tilde{\theta}_i \otimes_{*_n} \tilde{\phi}_i, \tag{4.8}$$

where $\tilde{a}_1, \tilde{a}_2, \ldots, \tilde{a}_{\min\{m,n\}}$ are the diagonal entries of S, $\tilde{b}_1, \tilde{b}_2, \ldots, \tilde{b}_{\min\{p,n\}}$ are the diagonal entries of M, and

$$\tilde{\phi}_j = \sum_{i=1}^{n} W_{ij} x_i,$$
$$\tilde{\psi}_j = \sum_{i=1}^{m} \tilde{U}_{ij} y_i,$$
$$\tilde{\theta}_j = \sum_{i=1}^{p} \tilde{V}_{ij} z_i.$$

The sets $\{\tilde{\phi}_1, \tilde{\phi}_2, \dots, \tilde{\phi}_n\}$, $\{\tilde{\psi}_1, \tilde{\psi}_2, \dots, \tilde{\psi}_n\}$, *and* $\{\tilde{\theta}_1, \tilde{\theta}_2, \dots, \tilde{\theta}_n\}$ *are orthonormal in* X_n, Y, *and* Z, *respectively. Here the* $*_n$*-inner product is used on* X_n.

Proof. Given that $H^{-1/2}A = USW^{-1}$, $J^{-1/2}B = VMW^{-1}$ is the GSVD of the matrix pair $H^{-1/2}A$, $J^{-1/2}B$, we know that $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{p\times p}$ are orthogonal. As noted above, the matrix representing $\hat{T}$ is $H^{-1}A$, and we have

$$H^{-1}A = (H^{-1/2}U)SW^{-1}.$$

For any $y = \sum_{i=1}^{m} \beta_i y_i$, $w = \sum_{i=1}^{m} \gamma_i y_i$, we have

$$\langle y, w\rangle_Y = \left\langle \sum_{j=1}^{m} \beta_j y_j, \sum_{i=1}^{m} \gamma_i y_i \right\rangle_Y = \sum_{i=1}^{m}\sum_{j=1}^{m} \langle y_j, y_i\rangle_Y \beta_j \gamma_i = \gamma \cdot H\beta.$$

It follows that

$$\left\langle \tilde{\psi}_j, \tilde{\psi}_i \right\rangle_Y = \left(\left(H^{-1/2}U\right)^T H \left(H^{-1/2}U\right)\right)_{ij} = (U^TU)_{ij} = \delta_{ij},$$

which shows that $\{\tilde{\psi}_1, \tilde{\psi}_2, \dots, \tilde{\psi}_n\}$ is orthonormal in Y. A similar proof shows that $\{\tilde{\theta}_1, \tilde{\theta}_2, \dots, \tilde{\theta}_n\}$ is orthonormal in Z. Also, we have

$$\begin{aligned} A^TH^{-1}A + B^TJ^{-1}B &= W^{-T}S^TU^TUSW^{-1} + W^{-T}M^TV^TVMW^{-1} \\ &= W^{-T}\left(S^TS + M^TM\right)W^{-1} \\ &= W^{-T}W^{-1}, \end{aligned}$$

which shows that

$$W^T\left(A^TH^{-1}A + B^TJ^{-1}B\right)W = I.$$

It follows that

$$\left\langle \sum_{k=1}^{n} W_{kj} x_k, \sum_{\ell=1}^{n} W_{\ell i} x_\ell \right\rangle_{*_n} = \left(W^T\left(A^TH^{-1}A + B^TJ^{-1}B\right)W\right)_{ij} = \delta_{ij}.$$

Therefore, $\{\tilde{\phi}_1, \tilde{\phi}_2, \dots, \tilde{\phi}_n\}$ is orthonormal with respect to the $*_n$-inner product.

Finally, it remains to prove that the equations (4.8) hold. We will verify the representation for $\hat{T}$; the proof for $\hat{L}$ is similar. Since $H^{-1}A$ is the matrix for $\hat{T}$ and $H^{-1}A = \tilde{U}SW^{-1}$, it suffices to show that

$$x = \sum_{j=1}^{n} \alpha_j x_j \;\Rightarrow\; \left(\sum_{i=1}^{\min\{m,n\}} \tilde{a}_i \tilde{\psi}_i \otimes_{*_n} \tilde{\phi}_i\right) x = \sum_{i=1}^{m} \beta_i y_i,$$

where

$$\beta = \tilde{U}SW^{-1}\alpha.$$

This is a direct calculation. We have

$$\left(\sum_{j=1}^{\min\{m,n\}} \tilde{a}_j \tilde{\psi}_j \otimes_{*_n} \tilde{\phi}_j\right) x = \sum_{j=1}^{n} \tilde{a}_j \left\langle \tilde{\phi}_j, x \right\rangle_{*_n} \tilde{\psi}_j,$$

$\tilde{\psi}_j = \sum_{i=1}^m \tilde{U}_{ij} y_i$, and

$$\begin{aligned}
\left\langle \tilde{\phi}_j, x \right\rangle_{*_n} = \sum_{k=1}^{n} \left\langle \tilde{\phi}_j, x_k \right\rangle_{*_n} \alpha_k &= \sum_{k=1}^{n} \left\langle \sum_{\ell=1}^{n} W_{\ell j} x_\ell, x_k \right\rangle_{*_n} \alpha_k \\
&= \sum_{k=1}^{n} \sum_{\ell=1}^{n} W_{\ell j} \left\langle x_\ell, x_k \right\rangle_{*_n} \alpha_k \\
&= \left(W^T \left(A^T H^{-1} A + B^T J^{-1} B\right) \alpha\right)_j \\
&= \left(W^{-1}\alpha\right)_j,
\end{aligned}$$

where we have used the fact that $W^T\left(A^T H^{-1} A + B^T J^{-1} B\right) W = I$. Therefore,

$$\begin{aligned}
\left(\sum_{j=1}^{\min\{m,n\}} \tilde{a}_j \tilde{\psi}_j \otimes_{*_n} \tilde{\phi}_j\right) x &= \sum_{j=1}^{\min\{m,n\}} \tilde{a}_j \left\langle \tilde{\phi}_j, x \right\rangle_{*_n} \tilde{\psi}_j \\
&= \sum_{j=1}^{\min\{m,n\}} \tilde{a}_j \left(W^{-1}\alpha\right)_j \left(\sum_{i=1}^{m} \tilde{U}_{ij} y_i\right) \\
&= \sum_{i=1}^{m} \left(\sum_{j=1}^{\min\{m,n\}} \tilde{a}_j \left(W^{-1}\alpha\right)_j \tilde{U}_{ij}\right) y_i \\
&= \sum_{i=1}^{m} \left(\tilde{U}SW^{-1}\alpha\right)_i y_i,
\end{aligned}$$

as desired. ∎

The scheme described by the above theorem does not fit into the standard framework of [2], and the convergence remains to be analyzed. Preliminary numerical experiments suggest that it is similar to that of the previous scheme. As noted above, this second approach has the advantage of avoiding the need to compute the matrix M (and also G), and it produces singular vectors $\{\tilde{\phi}_1, \tilde{\phi}_2, \ldots, \tilde{\phi}_n\}$, $\{\tilde{\psi}_1, \tilde{\psi}_2, \ldots, \tilde{\psi}_n\}$, and $\{\tilde{\theta}_1, \tilde{\theta}_2, \ldots, \tilde{\theta}_n\}$ that are orthonormal in X_n (under the $*_n$-inner product), Y, and Z, respectively.

4.3. Discussion. The special case discussed in this section is important and Theorem 4.2 is new, to the author's knowledge (a GSVE for a pair of operators was derived in the Ph.D. dissertation of Huang [8], but in that work both operators were required to be compact). However, one of the significant aspects of Theorem 2.2 is that it applies to noncompact operators. In particular, the operators T^*T and L^*L may have nonempty continuous spectra, which would imply that the representation of Theorem 4.2 is not suitable. Further work is needed to develop a computational scheme that is suitable when the continuous spectrum is nonempty.

REFERENCES

[1] A. Björck, *Least Squares Methods*, P. G. Ciarlet and J. L. Lions, eds., Handb. Numer. Anal., North-Holland, Amsterdam, 1990.
[2] D. Boffi, *Finite element approximation of eigenvalue problems*, Acta Numer., 19 (2010), pp. 1–120.
[3] G. B. Folland, *Real Anaysis: Modern Techniques and Their Applications*, Wiley, New York, 1984.
[4] M. S. Gockenbach, *Linear Inverse Problems and Tikhonov Regularization*, MAA Press, Washington, DC, 2016.
[5] G. H. Golub and C. F. V. Loan, *Matrix Computations*, 4th ed., Johns Hopkins University Press, Baltimore, 1996.
[6] P. R. Halmos, *What does the spectral theorem say?*, Amer. Math. Monthly, 70 (1963), pp. 241–247.
[7] P. C. Hansen, *Analysis of discrete ill-posed problems by means of the L-curve*, SIAM Rev., 34 (1992), pp. 561–580.
[8] Q. Huang, *Some Topics Concerning the Singular Value Decomposition and Generalized Singular Value Decomposition*, Ph.D. thesis, Arizona State University, Tempe, 2012.
[9] W. G. Kolata, *Approximation in variationally posed eigenvalue problems*, Numer. Math., 29 (1978), pp. 159–171.
[10] P. Lax, *Functional Analysis*, Wiley-Interscience, New York, 2002.
[11] J. Locker and P. M. Prenter, *Regularization with differential operators*. I. *General theory*, J. Math. Anal. Appl., 74 (1980), pp. 503–529.
[12] L. N. Trefethen and D. Bau, *Numerical Linear Algebra*, SIAM, Philadelphia, 1997.
[13] C. F. van Loan, *Generalizing the singular value decomposition*, SIAM J. Numer. Anal., 13 (1976), pp. 76–83.