Facultad de Ingeniería de la Universidad de Buenos Aires

Problemas de cuadrados mínimos con pesos en las clases de Schatten p y aplicaciones

Tesis doctoral presentada para optar por el título de doctorado de la Universidad de Buenos Aires

Maximiliano Contino

Directora
Dra. Alejandra Maestripieri

Co - Director
Dr. Juan Ignacio Giribet

Noviembre 2016

Maximiliano Contino: *Problemas de cuadrados mínimos con pesos en las clases de Schatten p y aplicaciones,* Tesis doctoral presentada para optar por el título de Doctor de la Universidad de Buenos Aires, © Fecha de entrega: Noviembre 2016

ÍNDICE

1 INTRODUCCIÓN

Sea $\mathcal{H}$ un espacio de Hilbert complejo y separable y $L(\mathcal{H})$ el álgebra de operadores lineales y acotados en $\mathcal{H}$. Dado un operador $A \in L(\mathcal{H})$, y un vector $b \in \mathcal{H}$, el método de cuadrados mínimos es un método clásico para obtener soluciones aproximadas, con algún criterio de optimalidad, de la ecuación

$$Ax = b,$$

cuando ésta no admite solución. Más precisamente, buscamos los vectores $x_0 \in \mathcal{H}$ tales que

$$\|Ax_0 - b\| = \min_{x\in\mathcal{H}} \|Ax - b\|,$$

donde $\|\cdot\|$ denota la norma asociada al producto interno de $\mathcal{H}$. Más aún, deseamos hallar en el conjunto de estas soluciones, aquellas que tengan norma mínima. Cuando el operador A tiene rango cerrado, entonces para todo $b \in \mathcal{H}$ existe una única solución de norma mínima. En este caso, el operador que asigna a cada $b \in \mathcal{H}$ la única solución del problema de cuadrados mínimos es lineal y acotado y es conocido como la inversa de Moore-Penrose de A y se denota $A^\dagger$. Los trabajos de Nashed y Votruba [68] y Nashed [67], y los libros de Rao y Mitra [72] y Ben-Israel y Greville [8] son excelentes referencias para consultar varios de los resultados sobre inversas generalizadas.

En algunos casos es necesario considerar otro producto escalar (no necesariamente equivalente al original) en el espacio $\mathcal{H}$ y entonces aparece naturalmente un peso (no inversible) y el concepto de inversas generalizadas. La primera aparición de inversas generalizadas con pesos se debe a Greville [50], quien las utilizó en problemas de ajuste de curvas y superficies. Esta noción ha aparecido en numerosas oportunidades en forma natural en problemas de matemática, estadística e ingeniería. Para nombrar algunos de los numerosos trabajos sobre este tema, mencionaremos a Chipman [16], quien reintrodujo esta noción para problemas de regresión lineal; Watson [85], Zyskind [86] y Rao y Mitra [72], hallaron aplicaciones a Estadística. Aplicaciones a procesamiento de imágenes y varios resultados algorítmicos, que utilizan inversas generalizadas con pesos singulares se pueden encontrar en los trabajos de Censor, Gordon y Gordon [14], [13] y Censor y Elfving, [11], [12].

En [39] Elden expone con gran detalle resultados sobre inversas generalizadas con pesos singulares, para espacios de dimensión finita. En este caso la existencia está garantizada. Sin embargo, cuando el espacio tiene dimensión infinita, la existencia de inversas generalizadas equivale a una cierta compatibilidad entre el peso (dado por un operador en $L(\mathcal{H})$ positivo, en general no inversible) y el rango (cerrado) de A. La noción de compatibilidad entre un operador positivo y un subespacio cerrado de $\mathcal{H}$ ha sido estudiada en varios trabajos, [25], [24], [26]. En [22], aplicando propiedades de compatibilidad, se determina cuándo el problema de cuadrados mínimos con pesos singulares admite solución y se caracteriza el conjunto de soluciones mediante pseudoinversas oblicuas.

Asociados a los problemas anteriores, aparecen ciertos problemas de minimización para operadores, del tipo cuadrados mínimos, pero considerando en ellos distintas normas de operadores, como por ejemplo las normas de Schatten p (o de Schatten von Neumann), para $1 \leq p \leq \infty$, y más generalmente aquellas normas que son unitariamente invariantes. Estas normas aparecen en distintos contextos, en muchos problemas matriciales y de operadores, e involucran los valores singulares de la matriz dada y más generalmente, en el caso de dimensión infinita, los autovalores del módulo del operador (compacto). Por ejemplo, en [47], G. R. Goldstein y J. A. Goldstein estudiaron el problema de hallar las matrices X_0 tales que,

$$|||AX_0 - I||| = \underset{X \in L(\mathcal{H})}{min} |||AX - I|||,$$

donde A es una matriz dada, I es la matriz identidad y $||| \cdot |||$ es una norma unitariamente invariante (como por ejemplo las normas p). Aplicando propiedades de funciones convexas, los autores caracterizaron la solución de norma mínima. En este trabajo estudiaremos una generalización del problema anterior en espacios de Hilbert de dimensión infinita. Motivados por aplicaciones en Procesamiento de Señales e Imágenes, introduciremos además un peso positivo que modificará el producto interno del espacio de Hilbert original obteniendo un semi-producto interno con el cual, el nuevo espacio no necesariamente resultará completo.

Problema a estudiar

Dado W un operador positivo, tal que $W^{1/2}$ es de la clase Schatten p, con $1 \leq p < \infty$, consideremos la semi-norma de operadores asociada a W dada por

$$\|X\|_{p,W} = \|W^{1/2}X\|_p, \; X \in L(\mathcal{H}).$$

Se estudiará el siguiente problema de aproximación de operadores para el peso W: dado $A \in L(\mathcal{H})$ de rango cerrado, y $B \in L(\mathcal{H})$, calcular

$$\underset{X \in L(\mathcal{H})}{inf} \|AX - B\|_{p,W}.$$

Observemos que este problema generaliza los casos estudiados en [47], para matrices, y en [22] donde se caracterizaron las soluciones para problemas con pesos singulares, para vectores. Observemos que cuando $\mathcal{H}$ tiene dimensión infinita, el peso W siempre es singular ya que es un operador compacto. Estudiaremos si el ínfimo se alcanza y caracterizaremos el conjunto de soluciones de este problema. Luego estudiaremos si existen soluciones de norma mínima, en este caso, la norma considerada será una norma q con otro peso W_2, tal que $W_2^{1/2}$ es de la clase de Schatten q, con $1 \leq q < \infty$.

Observar que, en dimensión finita, cuando el peso es la identidad, y $A, B \in S_p$, (donde S_p denota la clase de Schatten p) recobramos el problema mencionado en el punto anterior. Para este caso se puede ver en [40] y [67] un análisis de la existencia de soluciones y otras propiedades.

Soluciones de cuadrados mínimo con peso de la ecuación $AX - B = 0$

Un problema de interés en el área de Señales y Procesamiento de Imágenes es encontrar modelos de bajas dimensiones que aproximen, en algún sentido, determinados datos [31, 42]. En particular, varios de estos problemas se pueden enunciar de la siguiente manera: dada una matriz $B \in \mathbb{C}^{n\times n}$ con $ran(B) \geq k$ (donde $ran(B)$ denota el rango de la matriz B ó el número de columnas linealmente independientes), para $k \in \mathbb{N}$ tal que $k < n$, hallar una matriz $Y_0 \in \mathbb{C}^{n\times n}$ con $ran(Y_0) = k$ tal que,

$$Y_0 = \underset{\{Y \in \mathbb{C}^{n\times n}:\ ran(Y) = k\}}{argmin} f(Y - B),$$

para alguna función de *costo* $f : \mathbb{C}^{n\times n} \to \mathbb{R}$. Debido a su dificultad, este problema es estudiado usualmente relajando la restricción del rango de Y, lo cual bajo ciertas condiciones, resulta ser una relajación exacta (ver Sección 7.1). Para esto, se usa la factorización de $Y = AX$, con $A \in \mathbb{C}^{n\times k}, X \in \mathbb{C}^{k\times n}$.

Supongamos que la función *costo* viene dada por la norma de Frobenius $\|\cdot\|_F$, ahora estamos interesados en el siguiente problema:

$$Y_0 = \underset{\{X \in \mathbb{C}^{k\times n},\ A \in \mathbb{C}^{n\times k}\}}{argmin} \|AX - B\|_F. \tag{1.1}$$

De hecho, supongamos que $A_0 \in \mathbb{C}^{n\times k}, X_0 \in \mathbb{C}^{k\times n}$ satisfacen

$$\|A_0X_0 - B\|_F = \min_{\{X \in \mathbb{C}^{k\times n},\ A \in \mathbb{C}^{n\times k}\}} \|AX - B\|_F, \tag{1.2}$$

entonces,

$$\|A_0X_0 - B\|_F = \min_{X \in \mathbb{C}^{k\times n}} \|A_0X - B\|_F. \tag{1.3}$$

Si un peso (positivo) se introduce en la ecuación (1.3), o si la norma de Frobenius se reemplaza por otra norma unitariamente invariante, se puede estudiar el mismo problema.

La primera parte del Capítulo 6, está dedicada a buscar una solución de la extensión del Problema (1.3) a espacios de Hilbert abstractos.

Más específicamente, estudiamos el siguiente problema de aproximación: dado a $A \in L(\mathcal{H})$ un operador de rango cerrado, $B \in L(\mathcal{H})$ y W un operador positivo, tal que $W^{1/2}$ es de la clase Schatten p, con $1 \leq p < \infty$, analizamos las condiciones para la existencia de

$$\min_{X \in L(\mathcal{H})} \|AX - B\|_{p,W}. \tag{1.4}$$

Para estudiar el Problema 1.4, primero introducimos el siguiente problema asociado: dados $W \in L(\mathcal{H})^+$, A de rango cerrado y $B \in L(\mathcal{H})$ analizamos,

$$\min_{X \in L(\mathcal{H})} (AX - I)^* W (AX - I), \tag{1.5}$$

con el orden inducido en $L(\mathcal{H})$ por el cono de operadores positivos.

En el Capítulo 7, se describen algunos ejemplos de estos problemas de minimización en la Teoría de Control y en el Procesamiento de Señales [36, 80].

La existencia del mínimo de $\|AX - B\|_p$ en espacios de Hilbert, fue estudiada en [60] usando técnicas de diferenciación y también en [40] y [67], donde se estableció una conexión entre las normas p-Schatten y el orden en $L(\mathcal{H})^+$ (el orden inducido por el cono de operadores positivos). A partir de esto, la resolución del Problema 1.5, resulta de gran interés porque es posible relacionar las soluciones de este problema, en el que se considera el orden de operadores dado por el cono de positivos, con las del problema de minimización 1.4. De todas maneras, la introducción de un peso $W \in L(\mathcal{H})^+$ juega un importante rol, ya que estamos introduciendo en $\mathcal{H}$ un semi-producto interno asociado a W para el cual $\mathcal{H}$ ya no es más un espacio de Hilbert, a menos que W sea inversible. En este caso, la existencia de una proyección ortogonal adecuada no está asegurada. De hecho, la existencia de una proyección W-ortogonal en $R(A)$ (el subespacio cerrado rango de A) depende de la relación entre W y $R(A)$, lo que da lugar al concepto de *compatibilidad*.

La noción de compatibilidad, definida en [24] y desarrollada luego en [20, 23, 26], tiene sus orígenes en el trabajo de Z. Pasternak-Winiarski [69]. En dicho trabajo el autor estudió, para un subespacio cerrado fijo $\mathcal{S}$, la analiticidad del mapa $W \to P_{W,\mathcal{S}}$ que asocia a cada operador positivo inversible W la proyección ortogonal en $\mathcal{S}$ bajo el (equivalente) producto interno $\langle x, y \rangle_W = \langle Wx, y \rangle$, para $x, y \in \mathcal{H}$.

La noción de compatibilidad aparece cuando W puede ser cualquier operador semidefinido positivo, no necesariamente inversible (e incluso, un operador lineal autoadjunto acotado). Más precisamente, W y $\mathcal{S}$ se dicen *compatibles* si existe una proyección (lineal y acotada) Q con rango $\mathcal{S}$, la cual satisface $WQ = Q^*W$. Si W es positivo e inversible, existe una única proyección en $\mathcal{S}$ que es W-autoadjunta [24]. En general, puede suceder que no exista tal Q o que haya un número infinito de ellas. De todas maneras, existe una condición de ángulos entre $\mathcal{S}^\perp$ y $\overline{W(\mathcal{S})}$ que determina la existencia de dichas proyecciones [35], ver Teorema 4.1.5. La existencia de tales proyecciones está relacionada con la existencia del mínimo del Problema 1.4.

Resultados obtenidos

Por simplicidad en la sección 6.1.1, estudiamos el Problema 1.4 cuando $B = I$. En primer lugar, probamos que el ínfimo del conjunto $\{(AX - I)^*W(AX - I) : X \in L(\mathcal{H})\}$ (donde el orden es el inducido por el cono de operadores positivos), siempre existe y además que

$$\inf_{X \in L(\mathcal{H})} (AX - I)^*W(AX - I) = W_{/R(A)}, \tag{1.6}$$

donde $W_{/R(A)}$ es el operador de shorted de W a $R(A)$ (definido en la Sección 5.1), ver Proposición 6.1.1. En el Teorema 6.1.2, probamos que la existencia de mínimo del conjunto anterior es equivalente a la compatibilidad del par $(W, R(A))$ y también caracterizamos los operadores que minimizan este problema como las soluciones de la ecuación normal

$$A^*W(AX - I) = 0,$$

ó equivalentemente como las W-inversas de A (Proposición 6.1.4).

Resuelto el problema del mínimo para el orden de operadores, hicimos uso de la Proposición 3.1.15, la cual dice que si $|||\cdot|||$ es una norma unitariamente invariante en un ideal no nulo $\mathcal{J}$ de $L(\mathcal{H})$, y $S, T \in \mathcal{J}$. Entonces,

$$\text{si } T^*T \leq S^*S \text{ entonces } |||T||| \leq |||S|||.$$

De esta manera si $W^{1/2}$ está en la clase p-Schatten, para algún $1 \leq p < \infty$, es fácil concluir que si el par $(W, R(A))$ es compatible, entonces el mínimo del conjunto $\{\|AX - I\|_{p,W} : X \in L(\mathcal{H})\}$ existe, simplemente porque el mínimo en el orden de operadores garantiza el mínimo en las normas p, por la proposición anterior. Para probar la recíproca se utilizaron técnicas de diferenciación que nos permitieron concluir que si existe mínimo del conjunto

$$\{\|AX - I\|_{p,W} : X \in L(\mathcal{H})\}$$

entonces el mínimo cumple con la ecuación normal

$$A^*W(AX - I) = 0,$$

y por lo tanto el par $(W, R(A))$ es compatible, ver Teorema 6.1.5. De esta manera se establece la equivalencia entre la compatibilidad del par $(W, R(A))$ y la existencia del mínimo del Problema 1.4. En este caso, el conjunto de soluciones de 1.4 también son las W-inversas de A.

En la sección 6.1.2, probamos resultados similares para un operador arbitrario $B \in L(\mathcal{H})$, donde la existencia de mínimo del conjunto $\{\|AX - B\|_{p,W} : X \in L(\mathcal{H})\}$, es equivalente a la condición de compatibilidad $R(B) \subseteq R(A) + W(R(A))^\perp$ (Teorema 6.1.11). En este caso, los minimizantes son las W-inversas de A en $R(B)$ (ver definición en la Sección 4.2.2).

En la sección 6.2, se caracterizaron entre todos los operadores que realizan el mínimo de 1.4, aquellos que tienen norma mínima, donde la norma considerada es una norma q con otro peso W_2, tal que $W_2^{1/2}$ es de la clase de Schatten q, con $1 \leq q < \infty$.

En el caso donde $R(B) \subseteq R(A) + R(A)^{\perp_W}$, que es cuando se garantiza la existencia de mínimo del conjunto $\{(AX - B)^*W(AX - B) : X \in L(\mathcal{H})\}$, consideramos el conjunto

$$M_B = \{X_0 \in L(\mathcal{H}) : (AX_0 - B)^*W(AX_0 - B) = \min_{X \in L(\mathcal{H})} (AX - B)^*W(AX - B)\},$$

y dado $W_2 \in L(\mathcal{H})^+$, tal que $W_2^{1/2} \in S_q$, para algún q con $1 \leq q < \infty$, analizamos la existencia de

$$\min_{X \in M_B} \|X\|_{q,W_2}. \tag{1.7}$$

En el Teorema 6.2.1, se prueba que un operador X_1 resuelve el Problema 1.7 si y sólo si $R(W_2X_1) \subseteq N(WA)^{\perp}$. En este caso, los minimizantes resultan ser las WW_2-inversas de A en $R(B)$ (definición en la sección 6.2).

Soluciones de cuadrados mínimos con peso de la ecuación $AXB - C = 0$

A partir de los resultados obtenidos para la solución de cuadrados mínimos con peso de la ecuación $AX - C = 0$ detallados en la sección anterior, se buscó estudiar el problema más general que consiste en hallar las soluciones de cuadrados mínimos con peso de la ecuación $AXB - C = 0$, en este caso tanto A y B son operadores acotados de rango cerrado y C (como antes lo era B) es un operador acotado arbitrario. Este nuevo problema, resulta equivalente al anterior en los casos en que el operador B sea inversible o inyectivo. Para los caso en que el $N(B)$ (núcleo de B) sea distinto del trivial, se obtendrán resultados que se diferencian del caso anterior en donde tomábamos a $B = I$. Observar que ahora, el operador C ocupa el rol que tenía el operador B en el Problema 1.4 Por otra parte, en este caso supondremos que $B \neq 0$, pues si $B = 0$ el problema resulta trivial.

Los resultados obtenidos en este caso más general, tendrán aplicaciones en el área de Procesamiento de Señales, más precisamente en los problemas de *sampling* ó *muestreo*. El *sampling* ó *muestreo* es un operación que convierte una señal continua (modelada como vector en un espacio de Hilbert $\mathcal{H}$ apropiado) en una discreta. Frecuentemente las muestras de una señal $f \in \mathcal{H}$ son representadas de la siguiente manera: dado un marco $\{v_n\}_{n\in\mathbb{N}} \subseteq \mathcal{H}$ (ver Sección 7.2.1) de un subespacio cerrado $\mathcal{S}$, denominado subespacio de muestreo, dichas muestras vienen dadas por $\{f_n\}_{n\in\mathbb{N}} = \{\langle f, v_n \rangle\}_{n\in\mathbb{N}} \in \ell^2(\mathbb{N})$.

Por otra parte, dadas las muestras $\{f_n\}_{n\in\mathbb{N}} \in \ell^2(\mathbb{N})$, la señal de reconstrucción $\hat{f}$ viene dada por $\hat{f} = \sum_{n\in\mathbb{N}} f_n w_n$, donde $\{w_n\}_{n\in\mathbb{N}}$ es un marco del subespacio cerrado $\mathcal{R}$, llamado el subespacio de reconstrucción.

Supongamos que, A y B son los *operadores de sintesis* correspondientes a los marcos $\{w_n\}_{n\in\mathbb{N}}$ y $\{v_n\}_{n\in\mathbb{N}}$, respectivamente, es decir $A, B : \ell^2(\mathbb{N}) \to \mathcal{H}$ son los operadores tales que, si $x = \{x_n\}_{n\in\mathbb{N}} \in \ell^2(\mathbb{N})$, $Ax = \sum_{n\in\mathbb{N}} x_n w_n$ y $Bx = \sum_{n\in\mathbb{N}} x_n v_n$, que son acotados dado que $\{v_n\}_{n\in\mathbb{N}}$ y $\{w_n\}_{n\in\mathbb{N}}$ son marcos. Observar que, si $f \in \mathcal{H}$ es una señal, las muestras correspondientes están dadas por $\{f_n\}_{n\in\mathbb{N}} = B^* f$. Por otra parte, dadas las muestras $\{f_n\}_{n\in\mathbb{N}}$, la señal reconstruída viene dada por $\hat{f} = A\{f_n\}_{n\in\mathbb{N}}$, ver [38], [82].

Si sólo conocemos las muestras de una señal $\{f_n\}_{n\in\mathbb{N}} \in \ell^2(\mathbb{N})$, en general no es posible recuperar la señal $f \in \mathcal{H}$, incluso si se aplica un filtro digital (un operador lineal acotado $X : \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$) a estas muestras. Pero, en algunos casos es posible encontrar una buena representación de la señal $f \in \mathcal{H}$, es decir, una señal recuperada $\hat{f} = AXB^* f$ que tiene buenas propiedades. Por ejemplo, en el esquema clásico de muestreo (donde los subespacios de muestreo y de reconstrucción coinciden), es posible reconstruir el mejor aproximante de la señal f, es decir, es posible encontrar X tal que $AXB^* = P_{\mathcal{S}}$ y entonces $\hat{f} = P_{\mathcal{S}} f$, donde $P_{\mathcal{S}}$ es la proyección ortogonal sobre $\mathcal{S} = \mathcal{R} = R(A)$. Otro ejemplo interesante, donde los subespacios de muestreo y de reconstrucción podrían no coincidir, es el llamado esquema de muestreo consistente, donde las muestras de la señal reconstruída $\hat{f}$ son iguales a las muestras de la señal original f, es decir $B^* \hat{f} = B^* f$, en este caso X es tal que $Q = AXB^*$ resulta una proyección oblicua.

Consecuentemente, la señal reconstruida $\hat{f}$ no necesariamente es una buena aproximación de f, dado que la distancia $\|f - \hat{f}\| = \|f - AXB^* f\|$ no está minimizada. Ahora, supongamos que queremos encontrar un filtro digital $X \in \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$ tal que $AXB^* f$ sea una buena aproximación de f en $R(A) = \mathcal{R}$, es decir queremos que AXB^* aproxime a $P_{\mathcal{R}}$ en algún sentido. Por ejemplo, quisiéramos encontrar $X_0 : \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$ un operador lineal y acotado tal que, para cada $f \in \mathcal{H}$

$$\|(AX_0B^* - P_{\mathcal{R}})f\| \leq \|(AXB^* - P_{\mathcal{R}})f\|,$$

para cada $X \in L(\mathcal{H})$. Esto significa que, estamos interesados en el siguiente problema,

$$\min_{X\in L(\ell^2(\mathbb{N}))} (AXB^* - P_{\mathcal{R}})^*(AXB^* - P_{\mathcal{R}}),$$

con el orden inducido en $L(\mathcal{H})$ por el cono de operadores positivos.

Alternativamente, podemos aproximar $P_{\mathcal{R}}$ en alguna norma de operadores conveniente. Por ejemplo, en el caso finito dimensional, es usual considerar la norma de Frobenius; el problema asociado es equivalente a estudiar la existencia de

$$\min_{X\in L(\ell^2(\mathbb{N}))} \|AXB^* - P_{\mathcal{R}}\|_F,$$

donde $\|\cdot\|_F$ es la norma de Frobenius.

En la segunda parte del Capítulo 6, estudiamos una extensión de estos problemas. Más especíificamete dados $A, B \in L(\mathcal{H})$ de rango cerrado, $C \in L(\mathcal{H})$, $W \in L(\mathcal{H})$ positivo, estudiamos la existencia de

$$\min_{X \in L(\mathcal{H})} (AXB - C)^* W (AXB - C), \tag{1.8}$$

con el orden inducido en $L(\mathcal{H})$ por el cono de operadores positivos.

Si $W \in L(\mathcal{H})^+$ también satisface que $W^{1/2} \in S_p$, la clase p-Schatten (para algún p con $1 \leq p < \infty$), estudiamos la existencia de

$$\min_{X \in L(\mathcal{H})} \|AXB - C\|_{p,W}. \tag{1.9}$$

La existencia de mínimo de $\|AXB - C\|_p$ en espacios de Hilbert (sin peso), con hipótesis adecuadas para garantizar que $AXB - C \in S_p$, fueron estudiadas en [63] y en [62] usando técnicas de diferenciación. En [15], una caracterización de los puntos críticos (y equivalentemente de los mínimos globales) del mapa $\|AXB - C\|_p^p$ fueron dadas.

Resultados obtenidos

En la sección 6.3.1, estudiamos el Problema 1.8. Probamos que si $N(B) \subseteq N(A^*WC)$ entonces existe ínfimo del conjunto $\{(AXB - C)^* W (AXB - C) : X \in L(\mathcal{H})\}$ (donde el orden es el inducido por el cono de operadores positivos) y además cumple que

$$\inf_{X \in L(\mathcal{H})} (AXB - C)^* W (AXB - C) = C^* W_{/R(A)} C,$$

ver Proposición 6.3.2.

En el Teorema 6.3.3, probamos que el mínimo del conjunto anterior existe si sólo si $N(B) \subseteq N(A^*WC)$ y $R(C) \subseteq R(A) + W(R(A))^\perp$. Más aún, probamos que un operador X_0 minimiza este problema, si y sólo si $X_0 B$ es una W-inversa de A en $R(C)$.

Utilizando la Proposición 3.1.15, si $W^{1/2}$ está en la clase p-Schatten, para algún $1 \leq p < \infty$, entonces es fácil probar que si $N(B) \subseteq N(A^*WC)$ y $R(C) \subseteq R(A) + W(R(A))^\perp$, el mínimo del conjunto $\{\|AXB - C\|_{p,W} : X \in L(\mathcal{H})\}$ existe.

En el Lema 6.3.7, dimos una caracterización de los puntos críticos (y equivalentemente de los mínimos globales) del mapa $\|AXB - C\|_{p,W}^p$, que es similar a la presentada por [15], con la introducción de un peso W tal que $W^{1/2}$ está en la clase p-Schatten, para algún $1 \leq p < \infty$.

En el caso en que $p = 2$ ó alternativamente $1 \leq p < \infty$ y $N(B) \subseteq N(A^*WC)$, se probó que la existencia de mínimo del conjunto anterior es equivalente a la existencia de solución de la ecuación normal $A^*W(AXB - C)B^* = 0$. Finalmente, dimos algunos ejemplos para mostrar que en general, la existencia del mínimo del conjunto anterior, no es equivalente a la existencia de solución de la ecuación normal presentada. Estos hechos contradicen el Teorema 4.1 de [63], por lo cual, los resultados obtenidos en este trabajo son nuevos también para matrices.

Los contenidos de la tesis son los siguientes.

En el Capítulo 2, se introducen notaciones y algunos resultados que se utilizarán a lo largo del trabajo. Se presenta el importante Teorema de Douglas y se define el concepto de ángulos entre subespacios como así también algunas propiedades importantes. Se realiza una breve introducción a la teoría de operadores no acotados en espacios de Hilbert, se definen las inversas generalizadas y se demuestran propiedades importantes de las mismas.

En el Capítulo 3, se introducen las clases de Schatten S_p, con $1 \leq p \leq \infty$. Se definen operadores en dichas clases y se enuncian propiedades relevantes de los mismos. Se hace especial hincapié en las clases de Schatten S_1 (operadores de traza) y se muestra que el espacio S_p con la norma $\| \cdot \|_p$ resulta un espacio vectorial normado completo. Por otra parte, se definen las normas unitariamente invariantes y se establecen propiedades que se usarán a lo largo de la tesis. Por último, se prueba que la norma asociada al espacio de Schatten S_p, para $1 < p < \infty$, es Fréchet diferenciable y se deduce la fórmula de dicha derivada. Para el caso $p = 1$, se prueba la existencia de la derivada ϕ-direccional para todo $X \in S_p$ y en la dirección de todo $Y \in S_p$ y se da la fórmula de la misma.

En el Capítulo 4, dado un peso positivo $W \in L(\mathcal{H})$ y un subespacio cerrado $\mathcal{S} \subseteq \mathcal{H}$, se introduce el concepto de proyección W-autoadjunta y el de compatibilidad del par $(W, \mathcal{S})$. Se enuncian propiedades que deben cumplir los pares compatibles, como así también se dan algunos ejemplos de pares compatibles y no compatibles. Finalmente, se enuncian resultados que relacionan el ángulo entre subespacios y el concepto de compatibilidad y se presentan las proyecciones W-contractivas junto con algunas propiedades. A continuación, dado un operador $A \in L(\mathcal{H})$ de rango cerrado y un peso W positivo, se define el concepto de solución de cuadrados mínimos con peso de la ecuación $Az = x$. Finalmente se presentan las W-inversas de A, definidas inicialmente por S.K. Mitra y C.R. Rao en [72] y extendidas en este trabajo a las W-inversas de A en el rango de un cierto operador $B \in L(\mathcal{H})$, como los operadores $X_0 \in L(\mathcal{H})$, tales que para cada $x \in \mathcal{H}$, $X_0 x$ es una solución de cuadrados mínimos con peso de la ecuación $Az = Bx$, para cada $x, z \in \mathcal{H}$.

En el Capítulo 5, dado un operador positivo $W \in L(\mathcal{H})$ y un subespacio cerrado $\mathcal{S} \subseteq \mathcal{H}$, se presenta la noción de operador de *shorted* de W a $\mathcal{S}$. Dicha noción fue introducida por M.G. Krein [55] y fue luego redescubierta por W.N.Anderson y G.E. Trapp [3], [4], quienes lo aplicaron a la teoría de redes eléctricas. En dicho capítulo se presentan propiedades relevantes de los operadores de shorted y qué características adquieren cuando el par $(W, \mathcal{S})$ resulta compatible.

En el Capítulo 6, se estudia y se resuelve el Problema 1.4 y el problema asociado 1.6. En primer lugar, se resuelven dichos problemas para el caso $B = I$ y luego para cualquier $B \in L(\mathcal{H})$ arbitrario. A continuación, se caracterizan entre todos los operadores que minimizan el Problema 1.4, aquellos de norma mínima. Por último, se estudia y resuelve el Problema 1.9 y el problema asociado 1.8 y finalmente en la Sección 6.4, se estudian aplicaciones de estos resultados del operador de shorted.

En el Capítulo 7, se estudian aplicaciones de los problemas resueltos en el Capítulo 6. Por ejemplo, se estudian los problemas de optimización de rango fijo, y los problemas de muestreo y reconstrucción de señales.

Hacia el final del capítulo, decidimos presentar algunos ejemplos de problemas de Procrusto, los cuales se diferencian de los problemas que se resolvieron en este trabajo, en que tienen una condición adicional (restricción) que consta en general en buscar el mínimo pidiendo, por ejemplo, que el mismo resulte un operador unitario, etc. Si bien en la tesis no presentamos resultados para este tipo de problemas, nos pareció interesante incluir algunos ejemplos de este tipo de problemas de minimización los cuales surgen naturalmente, por ejemplo en Química Cuántica ó en el área de Procesamientos de Señales. Uno de ellos, por ejemplo, es el proceso de ortogonalización de Löwdin, [58], [1], donde se busca hallar una base ortonormal que minimice la desviación cuadrática con respecto a una base dada no ortonormal. Otro problema de minimización de operadores consiste en, dada una matriz A (no necesariamente autoadjunta) que representa una aproximación del Hamiltoniano (ver sección 7.3.1.1), hallar su estado básico, o en términos matemáticos, el autovalor mínimo. Sin embargo A podría no ser autoadjunta. Entonces se busca la matriz autoadjunta más cercana a A, donde la cercanía se determina en función de la norma usada. En este tipo de problemas se agrega una restricción sobre el conjunto de minimizantes, [49], [61]. Por último, daremos un ejemplo de problema de Procrusto conmumente encontrado en la teoría de marcos y otro relacionado con el procesamiento y reconstrucción de señales.

Finalmente, en el Capítulo 8, se enumeran las conclusiones más relevantes que resultan de la tesis.

En los anexos A y B, se recopilan varias generalidades conocidas de la teoría de espacios de Hilbert y se detallan las demostraciones de las propiedades enunciadas en el Capítulo 3 sobre espacios de Schatten S_p, respectivamente.

2

PRELIMINARES

En este capítulo se introducen las notaciones y algunos resultados que se utilizarán a lo largo de la tesis. Los temas tratados en este capítulo son muy diversos y no necesariamente están vinculados, decidimos incluirlos para que el trabajo resulte autocontenido.

Dados dos espacios de Hilbert complejos y separables $\mathcal{H}$ y $\mathcal{K}$, se denotará con $L(\mathcal{H},\mathcal{K})$ al álgebra de operadores lineales acotados de $\mathcal{H}$ en $\mathcal{K}$, $L(\mathcal{H}) = L(\mathcal{H},\mathcal{H})$, $L(\mathcal{H})^+$ al cono de operadores (semidefinidos) positivos de $L(\mathcal{H})$, $GL(\mathcal{H})$ al grupo de operadores inversibles de $L(\mathcal{H})$, $GL(\mathcal{H})^+ = GL(\mathcal{H}) \cap L(\mathcal{H})^+$, $CR(\mathcal{H})$ los operadores con rango cerrado de $L(\mathcal{H})$.

Para $A, B \in L(\mathcal{H})^+$, escribimos $A \leq B$, si $B - A \in L(\mathcal{H})^+$.

Dado un operador $T \in L(\mathcal{H})$, se utilizará la notación $R(T)$ como el rango o imagen de T y $N(T)$ para su núcleo. Dado $\mathcal{R}$ un subespacio cerrado de $\mathcal{H}$ y $\mathcal{S}$ un subespacio cerrado de $\mathcal{K}$, se denotará $L(\mathcal{R},\mathcal{S})$ al subespacio $L(\mathcal{R},\mathcal{S}) = \{T \in L(\mathcal{H},\mathcal{K}) : R(T) \subseteq \mathcal{S}, \mathcal{R}^\perp \subseteq N(T)\}$. Dado $T \in L(\mathcal{H})$ y $\mathcal{S}$ un subespacio cerrado de $\mathcal{H}$, se utilizará la notación $T_{|\mathcal{S}}$ para la restricción del operador T al subespacio $\mathcal{S}$, i.e., $T_{|\mathcal{S}} : \mathcal{S} \to \mathcal{H}$ es tal que para todo $x \in \mathcal{S}$, $Tx = T_{|\mathcal{S}}x$.

Se utilizará $\mathcal{Q} = \mathcal{Q}(L(\mathcal{H}))$ para denotar al conjunto de proyecciones en $L(\mathcal{H})$ y con $\mathcal{P} = \mathcal{P}(L(\mathcal{H}))$ al subconjunto de $\mathcal{Q}$ de proyecciones ortogonales. Los elementos en $Q \in \mathcal{Q} \setminus \mathcal{P}$ se denominan proyecciones oblicuas.

Dados dos subespacios $\mathcal{S}$ y $\mathcal{T}$ de $\mathcal{H}$, $\mathcal{S} \dotplus \mathcal{T}$ será la suma directa de $\mathcal{S}$ y $\mathcal{T}$, $\mathcal{S} \oplus \mathcal{T}$ la suma directa ortogonal de ellos, y $\mathcal{S} \ominus \mathcal{T} = \mathcal{S} \cap (\mathcal{S} \cap \mathcal{T})^\perp$. Si $\mathcal{H} = \mathcal{S} \dotplus \mathcal{T}$, la proyección sobre $\mathcal{S}$ a lo largo de $\mathcal{T}$ se denotará $P_{\mathcal{S}//\mathcal{T}}$ y es la única proyección con $R(P_{\mathcal{S}//\mathcal{T}}) = \mathcal{S}$ y $N(P_{\mathcal{S}//\mathcal{T}}) = \mathcal{T}$. En particular, $P_{\mathcal{S}} = P_{\mathcal{S}//\mathcal{S}^\perp}$ es la proyección ortogonal sobre $\mathcal{S}$.

Dados $x, y \in \mathcal{H}$, definimos $x \otimes y \in L(\mathcal{H})$ como $(x \otimes y)z = \langle z, y \rangle\, x,\ z \in \mathcal{H}$. En B.1.1, se enuncian las propiedades más relevantes del operador $x \otimes y$.

2.1 GENERALIDADES SOBRE OPERADORES EN ESPACIOS DE HILBERT

A continuación se enuncian algunos resultados de teoría de operadores en espacios de Hilbert, los cuales serán utilizados a lo largo de la tesis.

2.1.1 *Factorización de rangos de operadores*

El siguiente resultado, debido a R. G. Douglas [37], caracteriza las inclusiones de rangos de operadores.

Teorema 2.1.1. *Dados los espacios de Hilbert $\mathcal{H}$, $\mathcal{K}_1$, $\mathcal{K}_2$ y los operadores $A \in L(\mathcal{K}_1, \mathcal{H})$ y $B \in L(\mathcal{K}_2, \mathcal{H})$, las siguientes condiciones son equivalentes:*

1. *la ecuación $AX = B$ tiene una solución en $L(\mathcal{K}_2, \mathcal{K}_1)$;*
2. *$R(B) \subseteq R(A)$;*
3. *existe $\lambda > 0$ tal que $BB^* \leq \lambda AA^*$.*

En este caso, existe un único $D \in L(\mathcal{K}_2, \mathcal{K}_1)$ tal que $AD = B$ y $R(D) \subseteq \overline{R(A^*)}$. Además $D \in L(\mathcal{H})$ satisface $N(D) = N(B)$ y

$$\|D\| = \inf\{\lambda > 0 : BB^* \leq \lambda AA^*\}.$$

El operador D se denomina solución reducida *de la ecuación $AX = B$.*

Demostración. 1. $\Rightarrow$ 2. es trivial.

1. $\Rightarrow$ 3. : Dado que $CC^* \leq \|C\|^2 I$, si $B = AC$ con $C \in L(\mathcal{K}_2, \mathcal{K}_1)$ entonces $BB^* = ACC^*A^* \leq \|C\|^2 AA^*$.

2. $\Rightarrow$ 1. : Sean $A \in L(\mathcal{K}_1, \mathcal{H})$ y $B \in L(\mathcal{K}_2, \mathcal{H})$ tales que $R(B) \subseteq R(A)$. Se define el operador $C : \mathcal{K}_2 \to \mathcal{K}_1$ de la siguiente manera: si $x \in \mathcal{K}_2$, se tiene que $Bx \in R(B) \subseteq R(A)$ y, como $A : N(A)^\perp \to R(A)$ es biyectivo, existe un único $y \in N(A)^\perp$ tal que $Ay = Bx$. Definiendo $Cx = y$ se tiene que C está bien definido y $B = AC$.

Dado que C está definido en todo $\mathcal{K}_2$, para probar que C es acotado es suficiente ver que C tiene gráfico cerrado, [77, Teorema 2.15]. Sea $\{(x_n, y_n)\}_{n\in\mathbb{N}}$ una sucesión con $y_n = Cx_n$, tal que $(x_n, y_n) \to (x, y)$. Por definición, $Bx_n = Ay_n$ para todo $n \in \mathbb{N}$. Como A y B son acotados, $\lim\limits_{n\to\infty} Ay_n = Ay$ y $\lim\limits_{n\to\infty} Bx_n = Bx$, con lo cual $Ay = Bx$. Además, como $N(A)^\perp$ es cerrado, se sigue que $y \in N(A)^\perp$ y por consiguiente $Cx = y$.

3. $\Rightarrow$ 1. : Suponiendo que $BB^* \leq \lambda AA^*$ para algún $\lambda > 0$, se define la siguiente aplicación $D : R(A^*) \to R(B^*)$ tal que $D(A^*x) = B^*x$ para todo $x \in \mathcal{H}$. Entonces D está bien definida pues, si $A^*x = A^*x'$ entonces $A^*(x - x') = 0$ y, como $BB^* \leq \lambda AA^*$, se sigue que $x - x' \in N(B^*)$ o equivalentemente $B^*x = B^*x'$. Para ver que D es acotado en $R(A^*)$:

$$\|D(A^*x)\|^2 = \|B^*x\|^2 = \langle BB^*x, x \rangle \leq \lambda^2 \langle AA^*x, x \rangle = \lambda^2 \|A^*x\|^2.$$

Por lo tanto, D puede ser extendido de manera única a $\overline{R(A^*)}$. Definiendo $Dx = 0$ para $x \in R(A^*)^\perp$, entonces $D \in L(\mathcal{K}_1, \mathcal{K}_2)$ y $DB^* = A^*$, o equivalentemente, $A = BD^*$ con $D^* \in L(\mathcal{K}_2, \mathcal{K}_1)$.

Finalmente, para probar que existe una única solución reducida de la ecuación $AX = B$, observar que la inclusión $R(D) \subseteq \overline{R(A^*)} = N(A)^\perp$ identifica unívocamente al operador D, ya que $AD = B$ y $A_{|N(A)^\perp}$ es inyectivo. En particular, los operadores construidos en 2. $\Rightarrow$ 1. y 3. $\Rightarrow$ 1. tienen el rango contenido en el subespacio $\overline{R(A^*)} = N(A)^\perp$ y por lo tanto coinciden. $\square$

Corolario 2.1.2. *Sea $A \in L(\mathcal{H}, \mathcal{K})$, entonces $R((AA^*)^{1/2}) = R(A)$.*

Demostración. Dado que $AA^* = (AA^*)^{1/2}((AA^*)^{1/2})^*$, la igualdad entre los rangos de A y $(AA^*)^{1/2}$ es consecuencia del Teorema de Douglas. □

Corolario 2.1.3. *Sea $A \in L(\mathcal{H})$ y $Q \in \mathcal{Q}$ tal que $R(QA) \subseteq R(A)$. Entonces la solución reducida $D \in L(\mathcal{H})$ de la ecuación $AX = QA$ es una proyección, es decir, $D \in \mathcal{Q}$.*

Demostración. En primer lugar, observar que, si D es la solución reducida de $AX = QA$ entonces $AD^2 = QAD = Q^2A = QA$ y $R(D^2) \subseteq R(D) \subseteq \overline{R(A^*)}$. Por lo que, D^2 también es una solución reducida de la ecuación $AX = QA$ y, por la unicidad de la misma, resulta que $D \in \mathcal{Q}$. □

2.1.2 *Ángulos entre subespacios*

La noción de ángulo entre dos subespacios cerrados de un espacio de Hilbert permite dar una interpretación geométrica de resultados analíticos o topológicos. Las siguientes definiciones dan dos nociones diferentes de ángulos entre subespacios.

Definición (J. Dixmier [34]). El *ángulo mínimo* entre $\mathcal{S}$ y $\mathcal{T}$ es el ángulo $\alpha_0(\mathcal{S}, \mathcal{T}) \in [0, \frac{\pi}{2}]$ cuyo coseno está definido por

$$c_0(\mathcal{S}, \mathcal{T}) = \sup \{ |\langle x, y \rangle| : x \in \mathcal{S}, \|x\| \leq 1, \ y \in \mathcal{T}, \|y\| \leq 1 \}.$$

La idea de ángulo entre dos subespacios de dimensión uno es bastante natural. El ángulo de Dixmier entre $\mathcal{S}$ y $\mathcal{R}$ consiste en comparar los ángulos de todos los subespacios contenidos en $\mathcal{S}$ con dimensión uno y todos los subespacios de $\mathcal{R}$ con dimensión uno y tomar el ínfimo de estos ángulos. Cuando la intersección entre $\mathcal{S}$ y $\mathcal{R}$ es distinta de cero, esta noción de ángulo no es tan rica puesto que, independientemente de la posición relativa entre $\mathcal{S}$ y $\mathcal{R}$, el ángulo entre ellos es cero, basta tomar el subespacio generado por $x \in \mathcal{S} \cap \mathcal{R}$. En estos casos la siguiente noción de ángulo resulta más conveniente.

Definición (K. Friedrichs [45]). Dados dos subespacios cerrados $\mathcal{S}$ y $\mathcal{T}$ de $\mathcal{H}$, el *ángulo* entre ellos es aquel $\alpha(\mathcal{S}, \mathcal{T}) \in [0, \frac{\pi}{2}]$ cuyo coseno es

$$c(\mathcal{S}, \mathcal{T}) = \sup \{ |\langle x, y \rangle| : x \in \mathcal{S} \ominus \mathcal{T}, \|x\| \leq 1, \ \ y \in \mathcal{T} \ominus \mathcal{S}, \|y\| \leq 1 \}.$$

En general $c(\mathcal{S}, \mathcal{T}) \leq c_0(\mathcal{S}, \mathcal{T})$. Sin embargo, cuando $\mathcal{S} \cap \mathcal{T} = \{0\}$, estas dos definiciones coinciden.

El siguiente es un importate resultado que relaciona el ángulo entre dos subespacios cerrados $\mathcal{S}$, $\mathcal{T}$ y cuándo la suma $\mathcal{S} + \mathcal{T}$ es un conjunto cerrado.

Proposición 2.1.4. *Sean $\mathcal{S}$, $\mathcal{T}$ subespacios cerrados de $\mathcal{H}$. Las siguientes condiciones son equivalentes:*

1. $c(\mathcal{S}, \mathcal{T}) < 1$,
2. $\mathcal{S} + \mathcal{T}$ *es cerrado,*
3. $\mathcal{S}^{\perp} + \mathcal{T}^{\perp}$ *es cerrado,*
4. $c(\mathcal{S}^{\perp}, \mathcal{T}^{\perp}) < 1$,
5. $P_{\mathcal{S}^{\perp}}(\mathcal{T})$ *es cerrado.*

Demostración. Ver [32, Teorema 13]. □

Los siguientes resultados muestran cuándo la composición de dos operadores de rango cerrado tiene rango cerrado y cuando el rango de un operador coincide con el rango de su raíz cuadrada.

Lema 2.1.5. *Sea $A \in L(\mathcal{H})^{+}$. Entonces*

1. $N(A) = N(A^{1/2})$,
2. $R(A) \subseteq R(A^{1/2}) \subseteq \overline{R(A)}$,
3. *si $R(A)$ no es cerrado, entonces $R(A)$ está propiamente incluido en $R(A^{1/2})$.*

Demostración. Los ítems 1. y 2. son fáciles de ver. De hecho siempre vale que $N(A^{1/2}) \subseteq N(A)$ y por otra parte si $x \in N(A)$ entonces, $0 = \langle Ax, x \rangle = \langle A^{1/2}x, A^{1/2}x \rangle$ y tenemos la otra inclusión. Por otra parte, siempre tenemos que $R(A) \subseteq R(A^{1/2})$ y la segunda inclusión se deduce del ítem 1.

El ítem 3. lo probaremos por el contrarrecíproco. Si $R(A) = R(A^{1/2})$ y $\xi \in N(A)^{\perp} = N(A^{1/2})^{\perp} = \overline{R(A^{1/2})}$, entonces existe $\rho \in N(A)^{\perp}$ tal que $A^{1/2}\xi = A\rho$. Por lo tanto $A^{1/2}\rho - \xi \in N(A^{1/2}) \cap N(A^{1/2})^{\perp}$ y entonces $A^{1/2}\rho = \xi$ y $R(A^{1/2})$ es cerrado. Claramente, esto implica que $R(A)$ también es cerrado. □

Proposición 2.1.6. *Sea $\mathcal{H}$ un espacio de Hilbert y $A, B \in L(\mathcal{H})$ dos operadores con rango cerrado. Entonces, AB tienen rango cerrado si y sólo si $c(R(B), N(A)) < 1$.*

Demostración. Ver [9, 53]. □

2.1.3 *Operadores no acotados en espacios de Hilbert*

En esta sección, queremos analizar los operadores líneales en general, no necesariamente acotados. Estas definiciones y propiedades serán útiles porque no siempre lidiaremos con operadores acotados. Por ejemplo, si tenemos un operador $A \in L(\mathcal{H})$ cuyo rango no es cerrado, entonces el operador $A^\dagger$ (la inversa de Moore-Penrose de A), será un operador no acotado, ver [67]. Dado un peso $W \in L(\mathcal{H})^+$ y un subespacio cerrado $\mathcal{S} \subseteq \mathcal{H}$, otro ejemplo de operadores no acotados, los encontramos en las proyecciones W-simétricas de los pares cuasi-compatibles que definiremos en 4.1.

Definición. Sea $\mathcal{H}$ un espacio de Hilbert. Por operador en $\mathcal{H}$, ahora nos referiremos a un operador lineal cuyo dominio $D(T)$ es un subespacio de $\mathcal{H}$ y cuyo rango $R(T) \subseteq \mathcal{H}$.

No suponemos que el operador T es acotado ó continuo. Por supuesto que si T es acotado entonces tenemos que $D(T) = \mathcal{H}$.

Definiremos el *gráfico* de un operador T en $\mathcal{H}$ como el siguiente subespacio de $\mathcal{H} \times \mathcal{H}$

$$gr(T) = \{(x, Tx) \in \mathcal{H} \times \mathcal{H} : x \in D(T)\}.$$

Diremos que un operador en $\mathcal{H}$ es *cerrado* si su gráfico es un subespacio cerrado de $\mathcal{H} \times \mathcal{H}$. Entonces, por el Teorema del gráfico cerrado [77, Teorema 2.15], tenemos que $T \in L(\mathcal{H})$ si y sólo si T es cerrado y $D(T) = \mathcal{H}$.

Dados dos operadores lineales S, T en $\mathcal{H}$, usaremos la siguiente notación

$$T \subset S,$$

para denotar que $D(T) \subset D(S)$ y que $Tx = Sx,$ para cada $x \in D(T)$.

Definición (Definición de operador adjunto)**.** Sea T un operador densamente definido, es decir, $\overline{D(T)} = \mathcal{H}$. Definiremos el dominio de T^* como

$$D(T^*) = \{x \in \mathcal{H} : \exists\, z = z_x \in \mathcal{H} : \langle\, x, Ty\,\rangle = \langle\, z, y\,\rangle\,,\ \text{para cada } y \in D(T)\}.$$

Entonces

$$T^*x := z,\ (x \in D(T^*)).$$

Observar, que al suponer que T es densamente definido, nos aseguramos que de existir tal z, es único y entonces T^* está bien definido.

Las operaciones algebraicas ordinarias, con operadores no acotados, deben ser tratadas con cuidado. A continuación, las definiciones naturales para los dominios de las sumas y productos:

$$D(S+T) = D(S) \cap D(T).$$

$$D(ST) = \{x \in D(T) : Tx \in D(S)\}.$$

Las propiedades asociativas, $(R+S)+T = R+(S+T)$ y $(RS)T = R(ST)$, siguen valiendo. Con respecto a las propiedades distributivas, una de ellas, $(R+S)T = RT+ST$, vale en la forma usual. Pero la otra valdrá únicamente en la forma $T(R+S) \supset TR+TS$, ya que puede pasar que $(R+S)x \in D(T)$, aún cuando Rx ó Sx no pertenezca a $D(T)$.

La multiplicación por un escalar se define de la siguiente manera: Si $\alpha = 0$ entonces $D(\alpha T) = \mathcal{H}$ y $\alpha T = 0$, si $\alpha \neq 0$ entonces $D(\alpha T) = D(T)$ y $(\alpha T)x = \alpha(Tx)$, para $x \in D(T)$.

Teorema 2.1.7. *Sean S, T y ST operadores densamente definidos en $\mathcal{H}$. Entonces*

$$T^*S^* \subset (ST)^*.$$

Si además, $S \in L(\mathcal{H})$ entonces

$$T^*S^* = (ST)^*,$$

Demostración. Supongamos $x \in D(ST)$ e $y \in D(T^*S^*)$. Entonces

$$\langle Tx, S^*y \rangle = \langle x, T^*S^*y \rangle,$$

pues, $x \in D(T)$ y $S^*y \in D(T^*)$, y

$$\langle STx, y \rangle = \langle Tx, S^*y \rangle,$$

pues $Tx \in D(S)$ y además $y \in D(S^*)$. Por lo tanto

$$\langle STx, y \rangle = \langle x, T^*S^*y \rangle.$$

Esto prueba que

$$T^*S^* \subset (ST)^*.$$

Ahora, supongamos que $S \in L(\mathcal{H})$ e $y \in D((ST)^*)$. Entonces $S^* \in L(\mathcal{H})$, y por lo tanto $D(S^*) = \mathcal{H}$, además

$$\langle Tx, S^*y \rangle = \langle STx, y \rangle = \langle x, (ST)^*y \rangle,$$

para cada $x \in D(ST)$. Por lo tanto, $S^*y \in D(T^*)$, y entonces $y \in D(T^*S^*)$. Entonces usando la primera parte de la prueba, tenemos que

$$T^*S^* = (ST)^*.$$

□

Definición. Un operador lineal T en $\mathcal{H}$, densamente definido, se dice *simétrico* si

$$\langle Tx, y \rangle = \langle x, Ty \rangle, \text{ para cada } x, y \in D(T).$$

Por lo tanto, los operadores simétricos densamente definidos son exactamente aquellos que satisfacen

$$T \subset T^*.$$

Si $T = T^*$, entonces diremos que T es *autoadjunto*.

Estas dos propiedades, evidentemente coinciden cuando $T \in L(\mathcal{H})$. En general, no coinciden.

Más aún, si $D(T)$ es denso y $\langle Tx, y \rangle = \langle x, Sy \rangle$, para cada $x \in D(T)$ e $y \in D(S)$, entonces

$$S \subset T^*.$$

2.1.4 *Inversas generalizadas*

A continuación se enuncian algunos resultados sobre inversas generalizadas. Como se mostrará más adelante, las inversas generalizadas guardan una estrecha relación con el conjunto de proyecciones.

Sea $T \in L(\mathcal{H}, \mathcal{K})$, y sea $\mathcal{M} \subseteq \mathcal{H}$ un subespacio, tal que

$$\mathcal{H} = N(T) \dot{+} \mathcal{M}.$$

Sin suponer que $R(T)$ es un conjunto cerrado, sea $\mathcal{N}$ un subespacio tal que

$$\mathcal{K} = \overline{R(T)} \dot{+} \mathcal{N}.$$

Denotaremos $P = P_{N(T)//\mathcal{M}}$ y $Q = P_{\overline{R(T)}//\mathcal{N}}$, entonces es fácil ver que

$$T = T(I - P) = QT.$$

Observar que

$$T_{|\mathcal{M}} : \mathcal{M} \to R(T),$$

es un operador biyectivo. Entonces existe $(T_{|M})^{-1}$ y definimos

$$T^{\dagger}_{P,Q} y := (T_{|\mathcal{M}})^{-1} Q y, \tag{2.1}$$

$y \in R(T) \dot{+} \mathcal{N}$.

Es decir definimos la *inversa generalizada* $T^{\dagger}_{P,Q}$ relativa a dichas proyecciones, como la única extensión lineal de $(T_{|\mathcal{M}})^{-1}$ a $R(T) \dot{+} \mathcal{N}$, tal que $N(T^{\dagger}_{P,Q}) = \mathcal{N}$.

Si definimos el dominio de $T^{\dagger}_{P,Q}$ como

$$D(T^{\dagger}_{P,Q}) = R(T) \dot{+} \mathcal{N},$$

entonces $T^{\dagger}_{P,Q}$, resulta ser un operador cerrado densamente definido, que cumple las siguientes condiciones.

Teorema 2.1.8. *Las siguientes condiciones, caracterizan la inversa generalizada $T^{\dagger}_{P,Q}$ de un operador $T \in L(\mathcal{H}, \mathcal{K})$, con $D(T^{\dagger}_{P,Q}) = R(T) \dot{+} \mathcal{N}$.*

$$T^{\dagger}_{P,Q} T = I - P, \tag{2.2}$$

$$TT^{\dagger}_{P,Q} = Q_{|D(T^{\dagger}_{P,Q})}, \quad \textit{en } D(T^{\dagger}_{P,Q}), \tag{2.3}$$

$$T^{\dagger}_{P,Q} T T^{\dagger}_{P,Q} = T^{\dagger}_{P,Q}, \quad \textit{en } D(T^{\dagger}_{P,Q}), \tag{2.4}$$

$$TT^{\dagger}_{P,Q} T = T. \tag{2.5}$$

Demostración. Observar que para $y \in D(T^\dagger_{P,Q})$, vale que $Qy \in R(T)$, entonces $T^\dagger_{P,Q}$ está bien definido.

Observar que si $x \in \mathcal{H}$ lo descomponemos como $x = x_1 + x_2$, con $x_1 \in N(T)$ y $x_2 \in \mathcal{M}$, tenemos que

$$T^\dagger_{P,Q}Tx = (T_{|\mathcal{M}})^{-1}QTx = (T_{|\mathcal{M}})^{-1}Tx = (T_{|\mathcal{M}})^{-1}Tx_2 = x_2 = (I-P)x,$$

por lo tanto tenemos (2.2).

Además, como $R((T_{|\mathcal{M}})^{-1}Q) \subseteq \mathcal{M}$, entonces $TT^\dagger_{P,Q}x = T(T_{|\mathcal{M}})^{-1}Qx = Qx$, para $x \in D(T^\dagger_{P,Q})$, y obtenemos (2.3).

Por otra parte, sea $x \in D(T^\dagger_{P,Q})$, entonces por (2.2), tenemos que

$$T^\dagger_{P,Q}TT^\dagger_{P,Q}x = (I-P)T^\dagger_{P,Q}x = T^\dagger_{P,Q}x,$$

donde se usó que $R(T^\dagger_{P,Q}) = R((T_{|\mathcal{M}^\perp})^{-1}Q) \subseteq \mathcal{M}$, con lo que vale (2.4).

Finalmente, de (2.3), tenemos que para $x \in \mathcal{H}$ vale

$$TT^\dagger_{P,Q}Tx = Q_{|D(T^\dagger_{P,Q})}Tx = Tx,$$

y entonces obtenemos (2.5). □

Notar que $T^\dagger_{P,Q} \in L(\mathcal{H},\mathcal{K})$ si y sólo si $R(T)$ es cerrado, en este caso $D(T^\dagger_{P,Q}) = \mathcal{H}$.

Dado que los subespacios $N(T)$ y $\overline{R(T)}$ tienen infinitos complementos, T tendrá en general, infinitas inversas generalizadas.

De todas maneras, en los espacios de Hilbert, los complementos ortogonales de $N(T)$ y $\overline{R(T)}$, se distinguen entre los demás posibles complementos. En este caso, si $\mathcal{M} = N(T)^\perp$ y $\mathcal{N} = R(T)^\perp$ (y equivalentemente P y Q son proyecciones ortogonales), la inversa generalizada de T se llama la *inversa de Moore-Penrose* y se notará $T^\dagger$, notar que $D(T^\dagger) = R(T) \dotplus R(T)^\perp$. En este caso, las ecuaciones (2.4) y (2.5) y la condición que $TT^\dagger$ y $T^\dagger T$ sean simétricas, definen únicamente a $T^\dagger$. Observar que, $TT^\dagger$ no es autoadjunta, al menos que $R(T)$ sea cerrado.

Para el caso especial que $T \in L(\mathcal{H},\mathcal{K})$ y $R(T)$ es cerrado, $T^\dagger$ se puede caracterizar por las ecuaciones de Moore-Penrose:

$$(T^\dagger T)^* = T^\dagger T = P_{N(T)^\perp}, \tag{2.6}$$
$$(TT^\dagger)^* = TT^\dagger = P_{R(T)}, \tag{2.7}$$
$$T^\dagger TT^\dagger = T^\dagger, \tag{2.8}$$
$$TT^\dagger T = T. \tag{2.9}$$

Proposición 2.1.9. *Dado $T \in L(\mathcal{H},\mathcal{K})$, $R(T)$ es cerrado si y sólo si $R(T^*)$ es cerrado. Más aún, si R(T) es cerrado, entonces $(T^*)^\dagger = (T^\dagger)^*$.*

Demostración. Supongamos que $T \in L(\mathcal{H}, \mathcal{K})$ con $R(T)$ cerrado, entonces tenemos que $T^\dagger \in L(\mathcal{H})$ y se cumplen las ecuaciones (2.6), (2.7), (2.8), (2.9). Si calculamos el adjunto en dichas ecuaciones, tendremos que

$$T^\dagger T = T^*(T^\dagger)^*, \tag{2.10}$$
$$TT^\dagger = (T^\dagger)^* T^*, \tag{2.11}$$
$$(T^\dagger)^* T^* (T^\dagger)^* = (T^\dagger)^*, \tag{2.12}$$
$$T^*(T^\dagger)^* T^* = T^*. \tag{2.13}$$

Observar que, como $T^\dagger \in L(\mathcal{H})$ entonces por definición $(T^\dagger)^* \in L(\mathcal{H})$. Por lo tanto el operador $T^*(T^\dagger)^* \in L(\mathcal{H})$ y como cumple que

$$(T^*(T^\dagger)^*)^2 = T^*(T^\dagger)^* T^*(T^\dagger)^* = T^*(T^\dagger)^*,$$

entonces es una proyección. Además, tenemos que

$$R(T^*(T^\dagger)^*) \subseteq R(T^*) = R(T^*(T^\dagger)^* T^*) \subseteq R(T^*(T^\dagger)^*),$$

entonces

$$R(T^*(T^\dagger)^*) = N(I - T^*(T^\dagger)^*) = R(T^*),$$

y $R(T^*)$ es cerrado. La recíproca se prueba de manera similar.

La última afirmación resulta de ver que $(T^\dagger)^*$ satisface las condiciones de la definición de $(T^*)^\dagger$.

□

Observar que si $T \in L(\mathcal{H}, \mathcal{K})$ y $R(T)$ es cerrado, entonces

$$R(T^\dagger) = N(T)^\perp = R(T^*).$$

De hecho, $T^\dagger T = P_{N(T)^\perp}$, por lo que $N(T)^\perp \subseteq R(T^\dagger)$. Además $T^\dagger T T^\dagger = T^\dagger = P_{N(T)^\perp} T^\dagger$, por lo que $R(T^\dagger) \subseteq N(T)^\perp$. En forma similar, es fácil ver que $N(T^\dagger) = R(T)^\perp = N(T^*)$.

Proposición 2.1.10. *Sea $\mathcal{H}$ un espacio de Hilbert y $T \in L(\mathcal{H})$ de rango cerrado. Entonces,*

$$TT^* \text{ tiene rango cerrado y } R(T) = R(TT^*).$$

Demostración. Sea $T \in L(\mathcal{H})$ de rango cerrado, entonces $T^* \in L(\mathcal{H})$ también tiene rango cerrado. Por lo tanto, por la Proposición 2.1.6, como

$$c(R(T), N(T^*)) = c(R(T), R(T)^\perp) = 0 < 1,$$

vale que $R(TT^*)$ es cerrado. Entonces, por el Lema 2.1.5 y el Corolario 2.1.3, tenemos que

$$R(TT^*) = R((TT^*)^{1/2}) = R(T).$$

□

Definición. Dado $T \in L(\mathcal{H})$, se define *el módulo mínimo reducido* $\gamma(T)$ como

$$\gamma(T) = inf\{\|Tx\| : x \in N(T)^{\perp}, \|x\| = 1\}$$

Definición. Un operador $T \in L(\mathcal{H})$ se dice *acotado inferiormente* si existe un $\alpha > 0$ tal que $\|Tx\| \geq \alpha\|x\|$ para todo $x \in N(T)^{\perp}$.

Observar que un operador $T \in L(\mathcal{H})$ no nulo es acotado inferiormente si y sólo si $\gamma(T) > 0$.

Proposición 2.1.11. *Sea $T \in L(\mathcal{H})$ no nulo, entonces $\gamma(T) > 0$ si y sólo si $R(T)$ es cerrado.*

Demostración. Sea $T \in L(\mathcal{H})$ tal que $\gamma(T) > 0$. Si $y \in \overline{R(T)}$, existe una sucesión $\{x_n\}_{n\in\mathbb{N}}$ en $N(T)^{\perp}$ tal que $Tx_n \to y$. Luego, dado $\varepsilon > 0$, existe un $n_0 \in \mathbb{N}$ tal que $\|x_n\| \leq \frac{1}{\gamma(T)}\|Tx_n\| \leq \frac{1}{\gamma(T)}(\|y\| + \varepsilon)$ para todo $n \geq n_0$. Por lo tanto, existe $M > 0$ tal que $\|x_n\| \leq M$ para todo $n \in \mathbb{N}$.

Entonces, existe una subsucesión $\{x_{n_k}\}_{k\in\mathbb{N}}$ de $\{x_n\}_{n\in\mathbb{N}}$ y cierto vector $x_0 \in \overline{R(T)}$, tal que

$$\langle\, x_{n_k}, h \,\rangle \to \langle\, x_0, h \,\rangle\,,$$

para cada $h \in \mathcal{H}$. Es fácil ver que $\langle\, Tx_{n_k}, h \,\rangle \to \langle\, Tx_0, h \,\rangle$, para cada $h \in \mathcal{H}$ y además vale que $\langle\, Tx_{n_k}, h \,\rangle \to \langle\, y, h \,\rangle$, para cada $h \in \mathcal{H}$. Luego, por la unicidad del límite, resulta que $y = Tx_0$. Por lo tanto, $R(T)$ es cerrado.

Recíprocamente, si $R(T)$ es cerrado, $T^{\dagger}$ resulta acotada y dado $x \in N(T)^{\perp}$,

$$\|x\| = \|T^{\dagger}Tx\| \leq \|T^{\dagger}\|\|Tx\|.$$

Por lo tanto, $\gamma(T) \geq \|T^{\dagger}\|^{-1} > 0$. □

3 OPERADORES EN LAS CLASES DE SCHATTEN S_p

En este capítulo veremos ciertas clases S_p ($1 \leq p < \infty$), de operadores en un espacio de Hilbert $\mathcal{H}$, los cuales fueron introducidos por von Neumann y Schatten [79]. Cada una de estas clases es un ideal bilátero en $L(\mathcal{H})$ y consiste en operadores compactos. Por otra parte, cuando se define una norma acorde, S_p resulta un espacio de Banach (es decir un espacio vectorial normado sobre los complejos que resulta completo con la norma correspondiente), con propiedades análogas a aquellas que tienen las sucesiones del espacio $l_p(\mathbb{N})$, el cual se define como $l_p(\mathbb{N}) = \{\{x_n\}_{n\geq 1} \subseteq \mathbb{C} : \sum_{n\geq 1} |x_n|^p < \infty\}$, y está provisto de la norma $\|\{x_n\}_{n\geq 1}\|_p = (\sum_{n\geq 1} |x_n|^p)^{1/p}$.

La teoría del espacio S_p, sin embargo, tiene dificultades que provienen de la no conmutatividad de S_p, lo cual no tiene paralelismo con $l_p(\mathbb{N})$.

En este capítulo, supondremos que el espacio de Hillbert $\mathcal{H}$ es un espacio separable. Entonces cada base ortonormal de $\mathcal{H}$ se puede expresar como $\{\phi_j : j \in \mathbb{N}\}$, y un sistema ortonormal general (no necesariamente completo) lo notaremos como $\{\psi_k : k \in \mathbb{N}\}$.

Los resultados enunciados en este capítulo pueden ser ampliados en [75].

Definición. Cuando $1 \leq p < \infty$, S_p es el conjunto de todos los operadores $T \in L(\mathcal{H})$ compactos (ver A.2), que satisfacen la siguiente condición: para todo sistema ortonormal $\{\psi_k : k \in \mathbb{N}\}$ en $\mathcal{H}$,

$$\sum_{k\geq 1} |\langle T\psi_k, \psi_k \rangle|^p < \infty.$$

Usamos la notación $S_\infty = K(\mathcal{H}) = \{T \in L(\mathcal{H}) : T \text{ es compacto}\}$.

Observación 3.0.1. De la definición anterior, se desprende que S_p es un subespacio lineal de $L(\mathcal{H})$, y que si $T \in S_p$ entonces $T^* \in S_p$. De las propiedades del espacio $l_p(\mathbb{N})$, también se puede ver que $S_p \subseteq S_q$ si $1 \leq p \leq q \leq \infty$.

3.1 OPERADORES EN LA CLASE DE SCHATTEN

Lema 3.1.1. *Supongamos $1 \leq p < \infty$, T es un operador compacto y autoadjunto en $\mathcal{H}$, y $\{\lambda_n\}_{n\geq 1}$ es la sucesión de autovalores no nulos de T, contados de acuerdo a su multiplicidad. Entonces:*

1. *Si $T \in S_p$ entonces $\sum_{n\geq 1} |\lambda_n|^p < \infty$,*
2. *si $\sum_{n\geq 1} |\lambda_n|^p < \infty$, entonces $T \in S_p$ y para cualquier sistema ortonormal $\{\psi_k : k \in \mathbb{N}\}$ en $\mathcal{H}$,*

$$\sum_{k\geq 1} |\langle T\psi_k, \psi_k \rangle|^p \leq \sum_{n\geq 1} |\lambda_n|^p.$$

Demostración. Ver Anexo B, B.1.1. □

Es bien sabido que un operador T en $\mathcal{H}$ es compacto si y sólo si es el límite en norma de operadores de rango finito, ver Teorema A.2.1. El próximo teorema muestra un resultado similar para el espacio S_p.

Teorema 3.1.2. *Sea $T \in L(\mathcal{H})$ y $1 \leq p < \infty$. Entonces $T \in S_p$ si y sólo si existe una sucesión $\{F_n\}_{n\geq 1}$ de operadores en $\mathcal{H}$, tal que F_n tiene rango finito, no mayor que n y*

$$\sum_{n\geq 1} \|T - F_n\|^p < \infty.$$

Demostración. Ver B.1.2. □

Corolario 3.1.3. *$S_p(1 \leq p \leq \infty)$ contiene todo operador de rango finito de $\mathcal{H}$.*

Demostración. Se sigue del Teorema 3.1.2. □

Lema 3.1.4. *Supongamos que $1 \leq p \leq \infty$. Sea $T \in S_p$ y $A \in L(\mathcal{H})$. Entonces $AT \in S_p$ y $TA \in S_p$.*

Demostración. Para $p = \infty$ el resultado es evidente. Si $1 \leq p < \infty$, por el Teorema 3.1.2, existe una sucesión $\{F_n\}_{n\geq 1}$ de operadores en $\mathcal{H}$ de rango finito no mayor que n tal que $\sum_{n\geq 1} \|T - F_n\|^p < \infty$. Ya que AF_n y $F_n A$ tienen rango finito no mayor que n y además

$$\sum_{n\geq 1} \|AT - AF_n\|^p \leq \|A\|^p \sum_{n\geq 1} \|T - F_n\|^p < \infty,$$

y

$$\sum_{n\geq 1} \|TA - F_n A\|^p \leq \|A\|^p \sum_{n\geq 1} \|T - F_n\|^p < \infty,$$

se sigue, nuevamente del Teorema 3.1.2, que $AT \in S_p$ y $TA \in S_p$. □

Concluimos la sección resumiendo, en un teorema, las propiedades algebraicas principales de S_p que se han establecido hasta ahora.

Teorema 3.1.5. *Supongamos $1 \leq p \leq \infty$. Entonces S_p es un ideal bilátero en $L(\mathcal{H})$. Contiene todos los operadores de rango finito de $\mathcal{H}$ y también contiene el adjunto de sus miembros. Si $1 \leq p < \infty$, entonces cada elemento de S_p es un operador compacto.*

3.1.1 *Operadores de traza*

Definición. El ideal S_1 en $L(\mathcal{H})$ es llamado la clase de operadores de traza en $\mathcal{H}$. Si $T \in S_1$ y $\{\phi_j : j \in \mathbb{N}\}$ es una base ortonormal de $\mathcal{H}$, entonces la traza de T, denotada por $tr(T)$, se define por

$$tr(T) = \sum_{j\geq 1} \left\langle T\phi_j, \phi_j \right\rangle.$$

Observación 3.1.6. Observar que como $T \in S_1$, la $tr(T)$ converge absolutamente. Además, $tr(T)$ depende sólo de T (y no de la elección de la base ortonormal), de hecho, sea $\{e_n : n \in \mathbb{N}\}$ otra base ortonormal de $\mathcal{H}$, entonces

$$Te_n = \sum_{j\geq 1} \langle Te_n, \phi_j \rangle \phi_j.$$

Por lo tanto

$$\sum_{n\geq 1} \langle Te_n, e_n \rangle = \sum_{n\geq 1}\sum_{j\geq 1} \langle \langle Te_n, \phi_j \rangle \phi_j, e_n \rangle =$$

$$= \sum_{n\geq 1}\sum_{j\geq 1} \langle Te_n, \phi_j \rangle \langle \phi_j, e_n \rangle = \sum_{j\geq 1}\sum_{n\geq 1} \langle \phi_j, \langle \phi_j, Te_n \rangle e_n \rangle =$$

$$= \sum_{j\geq 1} \left\langle \phi_j, \sum_{n\geq 1} \langle T^*\phi_j, e_n \rangle e_n \right\rangle = \sum_{j\geq 1} \langle \phi_j, T^*\phi_j \rangle = \sum_{j\geq 1} \langle T\phi_j, \phi_j \rangle ,$$

donde el orden de la suma se puede intercambiar, debido a la convergencia absoluta de la serie con doble índice.

Teorema 3.1.7. *Sean $S, T \in S_1$, $A \in L(\mathcal{H})$ y $\alpha, \beta \in \mathbb{C}$ entonces:*

1. $tr(\alpha S + \beta T) = \alpha tr(S) + \beta tr(T)$,
2. $tr(S^*) = \overline{tr(S)}$,
3. $tr(S) > 0$, *si* $S \geq 0$ *y* $S \neq 0$,
4. $tr(AS) = tr(SA)$.

Demostración. Sea $\{\phi_j : j \in \mathbb{N}\}$ una base ortonormal de $\mathcal{H}$, los ítems 1. y 2. son una consecuencia inmediata de la definición de la traza.

Supongamos que $S \geq 0$, entonces $tr(S) = \sum_{j\geq 1} \langle S\phi_j, \phi_j \rangle \geq 0$, ya que cada término de la suma es no negativo. Si $tr(S) = 0$ entonces para cada $j \in \mathbb{N}$, tenemos

$$0 = \langle S\phi_j, \phi_j \rangle = \|S^{1/2}\phi_j\|^2,$$

entonces $S^{1/2} = 0$ y $S = 0$. Esto prueba el ítem 3.

Para el ítem 4., usando el Lema A.1.7 y la linealidad de la traza, vamos a suponer que $A \in L(\mathcal{H})$ es un operador unitario. En este caso $\{A\phi_j : j \in \mathbb{N}\}$ es una base ortonormal de $\mathcal{H}$ entonces

$$tr(SA) = \sum_{j\geq 1} \langle SA\phi_j, \phi_j \rangle =$$

$$\sum_{j\geq 1} \langle SA\phi_j, A^*A\phi_j \rangle = \sum_{j\geq 1} \langle AS(A\phi_j), A\phi_j \rangle = tr(AS).$$

□

3.1.2 *El espacio de Banach S_p*

En esta sección introduciremos una norma en S_p, y se probará que con dicha norma, S_p es un espacio de Banach. Supongamos que T es un operador lineal compacto en $\mathcal{H}$, y sea $T = U|T|$ (ver Teorema A.1.6). Entonces se sigue del Teorema A.2.2, que existe una sucesión decreciente $\{\mu_n\}_{n\geq 1}$ de números reales positivos (los autovalores de $|T|$, contados de acuerdo a su multiplicidad) y sucesiones ortonormales $\{\phi_n\}_{n\geq 1}$, $\{\psi_n\}_{n\geq 1}$, tales que

$$|T|x = \sum_{n\geq 1} \mu_n \langle x, \phi_n \rangle \phi_n, \tag{3.1}$$

$$Tx = \sum_{n\geq 1} \mu_n \langle x, \phi_n \rangle \psi_n. \tag{3.2}$$

Las sucesiones $\{\mu_n\}_{n\geq 1}$, $\{\phi_n\}_{n\geq 1}$, $\{\psi_n\}_{n\geq 1}$, son nulas a partir de k términos si T tiene rango finito k, y son infinitas si T no tiene rango finito.

Dado $p \geq 1$, la función $f_p(t) = t^p$, es continua en el eje real no negativo, y por lo tanto, también lo es en el espectro del operador positivo $|T|$. El operador $f_p(|T|)$ se denotará por $|T|^p$. Entonces, por el Teorema A.1.3, sigue que $|T|^p$ es compacto y que

$$|T|^p x = \sum_{n\geq 1} \mu_n^p \langle x, \phi_n \rangle \phi_n. \tag{3.3}$$

Lema 3.1.8. *Sea $1 \leq q \leq p < \infty$. Entonces las siguientes condiciones son equivalentes:*

1. $T \in S_p$,
2. $|T| \in S_p$,
3. $|T|^{p/q} \in S_q$.

Demostración. Ver B.1.3. □

Observación 3.1.9. Del Lema 3.1.1, y la equivalencia de los ítems 1. y 2. del Lema 3.1.8, se sigue que un operador compacto $T \in L(\mathcal{H})$ pertenece a S_p si y sólo si la sucesión $\{\mu_n\}_{n\geq 1}$ de los autovalores no nulos de $|T|$ satisface

$$\sum_{n\geq 1} \mu_n^p < \infty.$$

Usaremos esto como la definición de S_p.

Definición. Sea $1 \leq p < \infty$ y $T \in S_p$. Entonces

$$\|T\|_p = (tr\ (|T|^p))^{1/p}.$$

Usaremos la convención de que $\|\cdot\|_\infty$ es la norma usual de operadores en $K(\mathcal{H})$. En el Teorema 3.1.12 probaremos que $\|\cdot\|_p$ es efectivamente una norma.

Observación 3.1.10. Notar primero que, si $1 \le p < \infty$, $T \in S_p$ y la ecuación (3.1) se cumple, luego $|T|^p$ está representado por la ecuación (3.3), entonces

$$tr\ (|T|^p) = \sum_{n\ge1} \langle\, (|T|^p\phi_j, \phi_j \,\rangle = \sum_{n\ge1} \langle\, \mu_n^p\phi_j, \phi_j \,\rangle = \sum_{n\ge1} \mu_n^p.$$

Por lo tanto

$$\|T\|_p = (\sum_{n\ge1} \mu_n^p)^{1/p}, \tag{3.4}$$

donde $\{\mu_n\}_{n\ge1}$ es la sucesión decreciente de autovalores positivos de $|T|$, contados de acuerdo a su multiplicidad.

Por otra parte, dado que $\|T\| = \|\ |T|\ \| = \mu_1$, ver Lema A.1.5, tenemos que

$$\|T\| \le \|T\|_p. \tag{3.5}$$

Por último si T es un operador compacto normal, y $\{\lambda_n\}_{n\ge1}$ es la sucesión de autovalores no nulos de T (ordenados de acuerdo a valor absoluto decreciente y contado de acuerdo a su multiplicidades) entonces existe una sucesión ortonormal $\{\phi_n\}_{n\ge1}$ tal que

$$Tx = \sum_{n\ge1} \lambda_n \langle\, x, \phi_n \,\rangle\, \phi_n.$$

Dado que, $T\phi_n = \lambda_n\phi_n$, vale que $T^*\phi_n = \overline{\lambda_n}\phi_n$, y entonces

$$T^*Tx = \sum_{n\ge1} |\lambda_n|^2 \langle\, x, \phi_n \,\rangle\, \phi_n,$$

entonces

$$|T| = (T^*T)^{1/2} = \sum_{n\ge1} |\lambda_n| \langle\, x, \phi_n \,\rangle\, \phi_n.$$

Entonces, la sucesión de autovalores no nulos de $|T|$ es $\{|\lambda_n|\}_{n\ge1}$ y por lo tanto $T \in S_p$ si y sólo si

$$\sum_{n\ge1} |\lambda_n|^p < \infty.$$

Cuando esto sucede,

$$\|T\|_p = (\sum_{n\ge1} |\lambda_n|^p)^{1/p}.$$

Lema 3.1.11. *Si $T \in S_1$, entonces*

$$|tr\ (T)| \le \|T\|_1.$$

Demostración. Supongamos que T y $|T|$ están representados por las ecuaciones (3.1) y (3.2) respectivamente. Entonces $\mu_n > 0$ y $tr(|T|) = \sum_{n\ge1} \mu_n = \|T\|_1$. Entonces

$$|tr\ (T)| = |\sum_{n\ge1} \langle\, T\phi_n, \phi_n \,\rangle\, | \le$$

$$\sum_{n\geq 1} |\langle T\phi_n, \phi_n \rangle| = \sum_{n\geq 1} \mu_n |\langle \psi_n, \phi_n \rangle| \leq \sum_{n\geq 1} \mu_n = \|T\|_1.$$

□

Sea $\mathcal{F}$ el conjunto de todos los operadores de rango finito de $\mathcal{H}$ en $\mathcal{H}$. El siguiente teorema afirma que S_p es un espacio normado completo y además que el conjunto $\mathcal{F}$ es denso en S_p.

Teorema 3.1.12. *Sea $1 \leq p < \infty$. Entonces $\|T\|_p$ es una norma en S_p, y con esta norma S_p es un espacio de Banach. Además, el conjunto $\mathcal{F}$ de los operadores de rango finito en $\mathcal{H}$ es un subespacio denso de S_p.*

Demostración. Ver B.1.1. □

3.1.3 *Normas unitariamente invariantes*

Definición. Una norma $|||\cdot|||$ en un ideal no nulo $\mathcal{J}$ de $L(\mathcal{H})$ se llama unitariamente invariante si

$$|||UTV||| = |||T|||,$$

para todo operador unitario $U, V \in L(\mathcal{H})$ y $T \in \mathcal{J}$.

Las normas de Schatten p $\|\cdot\|_p$ y $\|\cdot\|$ (la norma de operadores en $L(\mathcal{H})$) son ejemplos de normas unitariamente invariantes.

Lema 3.1.13. *Toda norma unitariamente invariante $|||\cdot|||$ en un ideal no nulo $\mathcal{J}$ de $L(\mathcal{H})$ es simétrica, es decir,*

$$|||T_1 S T_2||| \leq \|T_1\| \; |||S||| \; \|T_2\|,$$

para todo $T_1, T_2 \in L(\mathcal{H})$ y $S \in \mathcal{J}$.

Demostración. Sea $T \in L(\mathcal{H})$, $S \in \mathcal{J}$ y consideremos un número real $\alpha > 1$. Por [70, Teorema 1], existen operadores unitarios $U_1, \cdots, U_n \in L(\mathcal{H})$ tales que

$$\frac{T}{\alpha\|T\|} = \frac{U_1 + \cdots + U_n}{n}.$$

Entonces

$$|||TS||| = \alpha\|T\| \; |||\frac{T}{\alpha\|T\|}S||| =$$

$$= \alpha\|T\| \; |||\frac{U_1 + \cdots + U_n}{n}S||| \leq \alpha\|T\| \sum_{i=1}^{n} \frac{1}{n}|||U_i S||| =$$

$$= \alpha\|T\| \sum_{i=1}^{n} \frac{1}{n}|||S||| = \alpha\|T\| |||S|||,$$

como α es arbitrario, la desigualdad $|||TS||| \leq \|T\| \; |||S|||$ se sigue. De manera similar podemos probar que $|||ST||| \leq \|T\| \; |||S|||$ □

Corolario 3.1.14. *Sea $|||\cdot|||$ una norma unitariamente invariante en un ideal no nulo $\mathcal{J}$ de $L(\mathcal{H})$, y $S, T \in \mathcal{J}$. Entonces,*

1. $|||T||| = |||\,|T|\,|||$,
2. $|||T||| = |||T^*|||$.

Demostración. Sea $T = U|T|$, la descomposición polar de T, entonces

$$|T| = U^*T, \quad T^* = |T|U^* \text{ y } |T| = T^*U.$$

Por lo tanto, usando el Lema 3.1.13, tenemos

$$|||T||| = |||U|T|\,||| \leq \|U\|\;|||\,|T|\,||| \leq |||\,|T|\,|||,$$

y por otro lado

$$|||\,|T|\,||| = |||U^*T||| \leq \|U^*\|\;|||T||| \leq |||T|||,$$

y se sigue el ítem 1. De manera similar se sigue el ítem 2. □

La siguiente proposición será útil para obtener varios resultados en este trabajo. Con hipótesis adicionales, una proposición similar se puede encontrar en [40, Proposition 2.5].

Proposición 3.1.15. *Sea $|||\cdot|||$ una norma unitariamente invariante en un ideal no nulo $\mathcal{J}$ de $L(\mathcal{H})$, y $S, T \in \mathcal{J}$. Entonces,*

$$Si\ T^*T \leq S^*S \text{ entonces } |||T||| \leq |||S|||.$$

Demostración. Si $T^*T \leq S^*S$, por el Teorema 2.1.1, existe un operador R con $\|R\| \leq 1$ tal que $T^* = S^*R$, entonces usando el Lema 3.1.13 tenemos

$$|||T||| = |||T^*||| = |||S^*R||| \leq |||S^*|||\;\|R\| \leq |||S^*||| = |||S|||.$$

□

3.2 RESULTADOS SOBRE DIFERENCIACIÓN EN S_p

A continuación daremos algunas definiciones que se usarán a lo largo de esta sección.

Definición. Sea $(\mathcal{E}, \|\cdot\|)$ un espacio de Banach y $f : \mathcal{E} \to \mathbb{R}$. Diremos que f es Fréchet diferenciable en $x \in \mathcal{E}$, si existe $D_xf : \mathcal{E} \to \mathbb{R}$, un operador lineal y continuo, tal que

$$lim_{z\to 0}\frac{f(x+z) - f(x)}{\|z\|} = D_xf(z).$$

Definición. Sea $(\mathcal{E}, \|\cdot\|)$ un espacio de Banach y $f : \mathcal{E} \to \mathbb{R}$. Sea $t \in \mathbb{R}$, la derivada de Gâteaux de f en el punto $x \in \mathcal{E}$ en dirección $y \in \mathcal{E}$ está definida por

$$Df(x,y) = lim_{t\to 0}\frac{f(x+ty)-f(x)}{t}.$$

Observación 3.2.1. Si f es Fréchet diferenciable, entonces es fácil ver que es derivable Gâteaux, además sus derivadas de Fréchet y Gâteaux conciden. De hecho, sea $z \in \mathcal{E}$, entonces

$$Df(x,z) - D_x f(z) = lim_{t\to 0}\frac{f(x+tz)-f(x)}{t} - D_x f(z) =$$

$$= lim_{t\to 0}\frac{f(x+th)-f(x)-D_x f(tz)}{t} =$$

$$= lim_{t\to 0}\frac{f(x+tz)-f(x)-D_x f(tz)}{\|tz\|}\|z\| = 0,$$

pues $t \to 0$ implica que $tz \to 0$ y $\|tz\| \to 0$.

Definición. Sea $(\mathcal{E}, \|\cdot\|)$ un espacio de Banach y $f : \mathcal{E} \to \mathbb{R}$. Sea $\phi \in [0, 2\pi)$ y $h > 0$, entonces la derivada $\phi-$direccional de f en el punto $x \in \mathcal{E}$, en la dirección $y \in \mathcal{E}$ está definida por

$$D_\phi f(x,y) = lim_{h\to 0^+}\frac{f(x+he^{i\phi}y)-f(x)}{h}.$$

A continuación probaremos varias propiedades de la derivada $\phi-$direccional de una función f y pondremos especial atención al caso en que la función f es la norma asociada al espacio de Banach $\mathcal{E}$ considerado.

Proposición 3.2.2. *Sea $(\mathcal{E}, \|\cdot\|)$ un espacio de Banach, $f : \mathcal{E} \to \mathbb{R}$, definida por $f(x) = \|x\|$. Sea $D_\phi f(x,y)$, la derivada $\phi-$direccional de f en el punto $x \in \mathcal{E}$ y en la dirección $y \in \mathcal{E}$ entonces:*

1. *La función*
$$\alpha_{x,y}(h) = \|x + he^{i\phi}y\|$$
es convexa,
2. *$D_\phi f(x,y)$ es la derivada a derecha de la función $\alpha_{x,y}$ en el punto $h = 0$, y teniendo en cuenta* 1., *$D_\phi f(x,y)$ siempre existe.*

Demostración.
1. Sea $0 \le t \le 1$, y $h_1, h_2 \in \mathbb{R}$ entonces

$$\alpha_{x,y}(th_1 + (1-t)h_2) = \|x + (th_1 + (1-t)h_2)e^{i\phi}y\| =$$

$$= \|t(x + h_1 e^{i\phi}y) + (1-t)(x + h_2 e^{i\phi}y)\| \le t\alpha_{x,y}(h_1) + (1-t)\alpha_{x,y}(h_2),$$

por lo tanto, la función $\alpha_{x,y}(h)$ es convexa.

2. Observar que $D_\phi f(x,y) = lim_{h\to 0^+} \frac{f(x+he^{i\phi}y)-f(x)}{h} = lim_{h\to 0^+} \frac{\alpha_{x,y}(h)-\alpha_{x,y}(0)}{h}$, por lo tanto $D_\phi f(x,y)$ es la derivada a derecha de la función $\alpha_{x,y}$ en el punto $h=0$. Además, por 1., como $\alpha_{x,y}$ es una función convexa, entonces existen las derivadas a derecha y a izquierda [84, Teorema1.6]. □

Proposición 3.2.3. *Sea $(\mathcal{E}, \|\cdot\|)$ un espacio de Banach, $f:\mathcal{E}\to\mathbb{R}$, definida por $f(x)=\|x\|$. Sea $D_\phi f(x,y)$, la derivada $\phi-$direccional de f en el punto $x\in\mathcal{E}$ y en la dirección $y\in\mathcal{E}$ entonces:*

1. *$D_\phi f(x,y)$ es una funcional en $\mathcal{E}$ subaditiva en la segunda variable y positiva.*
2. *$D_\phi f(x,e^{i\theta}y) = D_{\phi+\theta} f(x,y)$,*
3. *$|D_\phi f(x,y)| \le \|y\|$.*

Demostración. 1. : Tenemos que

$$\|x+he^{i\phi}(y_1+y_2)\| \le \|\frac{x}{2}+he^{i\phi}y_1\| + \|\frac{x}{2}+he^{i\phi}y_2\|,$$

y tomando límite obtenemos

$$D_\phi f(x,y_1+y_2) = lim_{h\to 0^+}\frac{\|x+he^{i\phi}(y_1+y_2)\|-\|x\|}{h} \le$$

$$lim_{h\to 0^+}\frac{\|x+2he^{i\phi}y_1)\|+\|x+2he^{i\phi}y_2)\|-2\|x\|}{2h} = D_\phi f(x,y_1)+D_\phi f(x,y_2).$$

2. : Es inmediato de la definición.
3. : Es suficiente ver que

$$|\ \|x+he^{i\phi}y\|-\|x\|\ | \le \|x+he^{i\phi}y-x\| = h\|y\|.$$

□

El siguiente lema nos permite relacionar, el mínimo de la función f (si existe) con la derivada $\phi-$direccional de f en el punto donde alcanza su mínimo.

Lema 3.2.4. *Sea $(\mathcal{E}, \|\cdot\|)$ un espacio de Banach y $f:\mathcal{E}\to\mathbb{R}$, tal que f tiene una derivada $\phi-$direccional para cada $\phi\in[0,2\pi)$, en cada punto $x\in\mathcal{E}$ y en cada dirección $y\in\mathcal{E}$. Si f tiene un mínimo global en $x_0\in\mathcal{E}$, entonces*

$$\underset{0\le\phi<2\pi}{inf}\ (D_\phi f(x_0,y)) \ge 0, \text{ para todo } y\in\mathcal{E}, \tag{3.6}$$

además si $f(x)=\|x\|$, también vale la recíproca.

Demostración. Supongamos que f tiene un mínimo global en $x_0 \in \mathcal{E}$, es decir,

$$f(x) \geq f(x_0), \text{ para todo } x \in \mathcal{E}.$$

Sean $h > 0$, $\phi \in [0, 2\pi)$ e $y \in \mathcal{E}$ tomados arbitrariamente, entonces si $x = x_0 + he^{i\phi}y$, tenemos que

$$f(x_0 + he^{i\phi}y) - f(x_0) \geq 0, \text{ para todo } \phi \in [0, 2\pi), h > 0, y \in \mathcal{E},$$

esto implica que

$$\frac{f(x_0 + he^{i\phi}y) - f(x_0)}{h} \geq 0, \text{ para todo } \phi \in [0, 2\pi), h > 0, y \in \mathcal{E},$$

haciendo tender $h \to 0^+$ tenemos que

$$D_\phi f(x_0, y) = \lim_{h \to 0^+} \frac{f(x_0 + he^{i\phi}y) - f(x_0)}{h} \geq 0, \text{ para todo } \phi \in [0, 2\pi), y \in \mathcal{E}.$$

Entonces

$$\inf_{0 \leq \phi < 2\pi} (D_\phi f(x_0, y)) \geq 0, \text{ para todo } y \in \mathcal{E}.$$

Recíprocamente, supongamos que que $f(x) = \|x\|$, entonces

$$D_\phi f(x_0, e^{i(\pi - \phi)}x_0) = \lim_{h \to 0^+} \frac{f(x_0 + he^{i\pi}x_0) - f(x_0)}{h} =$$

$$= \lim_{h \to 0^+} \frac{\|x_0 - hx_0\| - \|x_0\|}{h} = \|x_0\| \lim_{h \to 0^+} \frac{|1 - h| - 1}{h} = -\|x_0\|.$$

Sea $y \in \mathcal{E}$ arbitrario y sea $\hat{y} = y + e^{i(\pi - \phi)}x_0$. Usando (3.6) y la propiedades probadas en la Proposición 3.2.3, tenemos que

$$f(x_0) = \|x_0\| \leq -D_\phi f(x_0, e^{i(\pi - \phi)}x_0) + D_\phi f(x_0, \hat{y}) \leq$$

$$-D_\phi f(x_0, e^{i(\pi - \phi)}x_0) + D_\phi f(x_0, y) + D_\phi f(x_0, e^{i(\pi - \phi)}x_0) =$$

$$= D_\phi f(x_0, y) \leq \|y\| = f(y),$$

y x_0 resulta un mínimo global de $f(x) = \|x\|$. □

El siguiente lema prueba que, dado $\mathcal{T} \subseteq S_p$, con $1 < p < \infty$, un conjunto convexo, de existir el mínimo del conjunto

$$\{\|X\|_p : X \in \mathcal{T}\},$$

este es único.

Lema 3.2.5. *Sea $\mathcal{T}$ un conjunto convexo de operadores en S_p, con $1 < p < \infty$, entonces existe a lo sumo un minimizante de $\|X\|_p$ si $X \in \mathcal{T}$.*

Demostración. Supongamos que existen dos minimizantes distintos de $\|X\|_p$, $A_1, A_2 \in \mathcal{T}$. Entonces

$$\|A_1\|_p = \|A_2\|_p \leq \|X\|_p, \text{ para cada } X \in \mathcal{T}.$$

Como $\mathcal{T}$ es convexo, tenemos que $\frac{A_1+A_2}{2} \in \mathcal{T}$. Entonces debe valer que

$$\|A_1 + A_2\|_p < \|A_1\|_p + \|A_2\|_p,$$

de lo contrario, si $\|A_1 + A_2\|_p = \|A_1\|_p + \|A_2\|_p$, para $1 < p < \infty$, por [64, Teorema 2.41], tendríamos que $a_1 A_1 = a_2 A_2$, para algunos números reales no negativos a_1, a_2 (con $a_1 + a_2 > 0$), y dado que $\|A_1\|_p = \|A_2\|_p$, tendríamos que $A_1 = A_2$.

Por lo tanto, para todo $X \in \mathcal{T}$, tenemos:

$$\|\frac{A_1 + A_2}{2}\|_p < \frac{\|A_1\|_p + \|A_2\|_p}{2} = \|A_1\|_p (= \|A_2\|_p) \leq \|X\|_p,$$

lo que contradice que A_1 (ó A_2) es un minimizante de $\|X\|_p$ en $\mathcal{T}$.

□

3.2.1 *Fórmula de derivadas en S_p*

En este sección probaremos, para el caso $1 < p < \infty$, que la norma asociada al espacio de Schatten correspondiente S_p, es Fréchet diferenciable y deduciremos la fórmula de dicha derivada. Para el caso $p = 1$, se demostrará que la norma $\|\cdot\|_1$ tiene derivada $\phi-$direccional y también se deducirá la fórmula de la misma. Debido a la importancia que estos resultados tienen para el trabajo, se decidió incluir demostraciones detalladas de los mismos. Estos resultados aparecieron originalmente en [2](caso $1 < p < \infty$) y en [54] (caso $p = 1$).

Caso $1 < p < \infty$

Para el caso $1 < p < \infty$, probaremos que el espacio $(S_p, \|\cdot\|_p)$, con $1 < p < \infty$, es un espacio uniformemente convexo. Está propiedad nos permite afirmar que $\|\cdot\|_p$, con $1 < p < \infty$, es Fréchet diferenciable y luego deduciremos la fórmula de dicha derivada.

Definición. Sea $(\mathcal{E}, \|\cdot\|)$ un espacio de Banach. Diremos que $\mathcal{E}$ es uniformemente convexo si para todo $\varepsilon > 0$, existe $\delta > 0$ tal que

$$\|x\| = \|y\| = 1, \text{ y } \|x + y\| > 2 - \delta \text{ implican } \|x - y\| < \varepsilon.$$

Lema 3.2.6. *$(S_p, \|\cdot\|_p)$, con $1 < p < \infty$, es un espacio uniformemente convexo.*

Demostración. Sean $S, T \in S_p$, con $1 < p < \infty$ y sea $\frac{1}{q} + \frac{1}{p} = 1$. C. A. MCCarthy demostró en [64, Teorema 2.7] las siguientes desigualdades:

si $1 < p \leq 2$ entonces

$$\|T+S\|_p^q + \|T-S\|_p^q \leq 2(\|T\|_p^p + \|S\|_p^p)^{\frac{q}{p}},$$

si $2 \leq p < \infty$ entonces

$$\|T+S\|_p^p + \|T-S\|_p^p \leq 2^{p-1}(\|T\|_p^p + \|S\|_p^p).$$

Si tomamos $\|S\|_p = \|T\|_p = 1$ en las desigualdades de arriba, tenemos

$$\|T-S\|_p^q \leq 2^q - \|T+S\|_p^q,\ 1 < p \leq 2,$$

y

$$\|T-S\|_p^p \leq 2^p - \|T+S\|_p^p,\ 2 \leq p < \infty.$$

Dado $2 > \varepsilon > 0$, para $1 < p \leq 2$, sea $\delta = 2 - (2^q - \varepsilon^q)^{1/q} > 0$, entonces si $\|S+T\|_p > 2 - \delta = (2^q - \varepsilon^q)^{1/q}$, tenemos que

$$\|T-S\|_p^q \leq 2^q - \|T+S\|_p^q < 2^q - 2^q + \varepsilon^q = \varepsilon^q,$$

entonces $\|T-S\|_p < \varepsilon$.

Dado $2 > \varepsilon > 0$, para $2 \leq p < \infty$, sea $\delta = 2 - (2^p - \varepsilon^p)^{1/p} > 0$, si $\|S+T\|_p > 2 - \delta = (2^p - \varepsilon^p)^{1/p}$, tenemos que

$$\|T-S\|_p^p \leq 2^p - \|T+S\|_p^p < 2^p - 2^p + \varepsilon^p = \varepsilon^p,$$

y entonces $\|T-S\|_p < \varepsilon$. Concluyendo que $(S_p, \|\cdot\|_p)$, con $1 < p < \infty$ es uniformemente convexo. □

Definición. Sean $\mathcal{H}, \mathcal{K}$ espacios de Hilbert y $T_n, T \in L(\mathcal{H}, \mathcal{K})$. Decimos que T_n converge fuertemente (o en sentido fuerte) a T si para cualquier $x \in \mathcal{H}$ se tiene que $T_n(x) \underset{n\to\infty}{\to} T(x)$.

Teorema 3.2.7. *Sea $1 < p < \infty$ y $G_p : S_p \to \mathbb{R}^+$, $G_p(X) = \|X\|_p^p$. Sean $X, Y \in S_p$, entonces G_p es Frechét diferenciable y su derivada en $X \in S_p$ está dada por*

$$D_X G_p(Y) = p\ Re\ [tr(|X|^{p-1}U^*Y)],$$

donde $Re(z)$ es la parte real del número complejo z, $tr(T)$ denota la traza del operador T y $X = U|X|$, es la descomposición polar del operador X, con U la isometría parcial tal que $N(U) = N(X)$.

Demostración. En el Lema 3.2.6, vimos que S_p $(1 < p < \infty)$ es uniformemente convexo. Se sigue entonces de un resultado estándar (ver, por ejemplo [33, Teorema 1, pág. 36]) que entonces S_p tiene una norma diferenciable Fréchet. Entonces G_p es diferenciable Fréchet y sólo resta establecer la fórmula de su derivada. Sean $X, Y \in S_p$ y sea E cualquier proyección que conmute con X. Entonces, afirmamos que

$$D_X G_p(EY(I-E)) = 0.$$

Un simple cálculo permite ver que

$$(2E-I)[X+EY(I-E)](2E-I) = X - EY(I-E). \tag{3.7}$$

Observar que $2E-I$ es unitario, es decir, $(2E-I)^*(2E-I) = (2E-I)(2E-I)^* = I$, entonces vale que

$$\|X+EY(I-E)\|_p^p = \|X-EY(I-E)\|_p^p,$$

ó en otras palabras, $\|X+EY(I-E)\|_p^p$ es una función par de Y. Por lo tanto, su derivada se anula en 0, lo cual permite concluir que $D_X G_p(EY(I-E)) = 0$.

Ahora, consideremos el caso en que X es positivo. Entonces por el Teorema A.2.2

$$X = \sum_{i \geq 1} \lambda_i x_i \otimes x_i,$$

donde $\lambda_i \geq 0$ y $\{x_i\}_{i\geq 1}$ es una base ortonormal. Sea E_i una proyección en el espacio generado por x_i y sea $F_n = I - \sum_{i=1}^n E_i$. Ya que, por lo de arriba, $D_X G_p(E_1 Y F_1) = D_X G_p(F_1 Y E_1) = 0$, tenemos

$$D_X G_p(Y) = D_X G_p(E_1 Y E_1) + D_X G_p(F_1 Y F_1).$$

Repitiendo el argumento de arriba, obtenemos por inducción, que para todo $n \in \mathbb{N}$,

$$D_X G_p(Y) = \sum_{i=1}^n D_X G_p(E_i Y E_i) + D_X G_p(F_n Y F_n). \tag{3.8}$$

Veamos que

$$D_X G_p(E_i Y E_i) = pRe\,[\lambda_i^{p-1} \langle Yx_i, x_i \rangle].$$

De hecho, notemos que

$$X + tE_i Y E_i = \sum_{j \geq 1} \lambda_j x_j \otimes x_j + t x_i \otimes x_i Y x_i \otimes x_i =$$

$$= \sum_{j \neq i} \lambda_j x_j \otimes x_j + \lambda_i x_i \otimes x_i + t x_i \otimes x_i Y x_i \otimes x_i =$$

$$= \sum_{j \neq i} \lambda_j x_j \otimes x_j + (\lambda_i + t \langle Yx_i, x_i \rangle) x_i \otimes x_i,$$

entonces

$$\|X + tE_i Y E_i\|_p^p = \sum_{j \neq i} \lambda_j^p + |\lambda_i + t \langle Yx_i, x_i \rangle|^p.$$

Llamando $\beta_i = \langle Yx_i, x_i \rangle$ tenemos que

$$\frac{d|\lambda_i + t\beta_i|^2}{dt}|_{t=0} = lim_{h\to 0}\frac{|\lambda_i + h\beta_i|^2 - \lambda_i^2}{h} =$$

$$= lim_{h\to 0}\frac{(\lambda_i + h\beta_i)(\lambda_i + h\overline{\beta_i}) - \lambda_i^2}{h} =$$

$$= lim_{h\to 0}\frac{\lambda_i^2 + 2hRe(\lambda_i\beta_i) + h^2|\beta_i|^2 - \lambda_i^2}{h} = 2Re(\lambda_i\beta_i).$$

Por lo tanto

$$\frac{d|\lambda_i + t\beta_i|^p}{dt}|_{t=0} = \frac{d(|\lambda_i + t\beta_i|^2)^{\frac{p}{2}})}{dt}|_{t=0} =$$

$$= \frac{p}{2}(|\lambda_i + t\beta_i|^2)^{\frac{p}{2}-1}|_{t=0}\frac{d|\lambda_i + t\beta_i|^2}{dt}|_{t=0} = pRe(\lambda_i^{p-1}\beta_i).$$

Entonces

$$D_XG_p(E_iYE_i) = lim_{t\to 0}\frac{\|X + tE_iYE_i\|_p^p - \|X\|_p^p}{t} =$$

$$= lim_{t\to 0}\frac{|\lambda_i + t\langle Yx_i, x_i\rangle|^p - \lambda_i^p}{t} = \frac{d|\lambda_i + t\beta_i|^p}{dt}|_{t=0} = pRe(\lambda_i^{p-1}\beta_i),$$

donde usamos el hecho de que, debido a que G_p es Fréchet diferenciable, su derivada de Fréchet y Gâuteaux coinciden.

Por lo tanto

$$D_XG_p(E_iYE_i) = pRe(\lambda_i^{p-1}\langle Yx_i, x_i\rangle) = pRe(\left\langle X^{p-1}Yx_i, x_i \right\rangle),$$

entonces volviendo a la ecuación (3.8) tenemos que

$$D_XG_p(Y) = p\sum_{i=1}^{n} Re(\left\langle X^{p-1}Yx_i, x_i \right\rangle) + D_XG_p(F_nYF_n).$$

Debido a que $\{F_n\}_{n\geq 1}$ converge fuertemente a 0 cuando $n \to \infty$, se sigue (ver [41, Teorema 1]) que F_nYF_n converge a 0 en S_p. Debido a que D_XG_p es continua (por definición), tenemos que $D_X(F_nYF_n) \to 0$ y entonces tomando límite en la ecuación anterior, tenemos que

$$D_XG_p(Y) = p\sum_{i=1}^{\infty} Re(\left\langle X^{p-1}Yx_i, x_i \right\rangle) = pRe\ tr(X^{p-1}Y),$$

que es la fórmula requerida cuando $X \geq 0$.

Ahora, sea $X \in S_p$ y sea $X = U|X|$ su descomposición polar. Es bien sabido que existe V tal que V ó V^* es una isometría tal que V y U coinciden en $N(|X|)^\perp$. Entonces $X = V|X|$.

Si V^* es una isometría, entonces para cualquier $Y \in S_p$ tenemos (usando que $VV^* = I$)

$$|X + Y|^2 = (|X|V^* + Y^*)(V|X| + Y) = |X|^2 + |X|V^*Y + Y^*V|X| + Y^*VV^*Y =$$

$$= ||X| + V^*Y|^2,$$

entonces

$$\|X+Y\|_p^p = \||X| + V^*Y\|_p^p,$$

y entonces $D_X G_p(Y) = D_{|X|}(V^*Y)$, y por lo tanto

$$D_X G_p(Y) = D_{|X|}(V^*Y) = pRe\ tr(|X|^{p-1}V^*Y) = pRe\ tr(|X|^{p-1}U^*Y),$$

y tenemos la fórmula requerida.

En caso en que V sea una isometría, tomando adjuntos y realizando cálculos similares tenemos que

$$D_X G_p(Y) = D_{|X^*|}(VY^*) = pRe\ tr(|X^*|^{p-1}VY^*) = pRe\ tr(V|X|^{p-1}V^*VY^*) =$$

$$= pRe\ tr(U|X|^{p-1}Y^*) = pRe\ tr(|X|^{p-1}Y^*U) = pRe\ tr((U^*Y|X|^{p-1})^*) =$$
$$= pRe\ tr(U^*Y|X|^{p-1}) = pRe\ tr(|X|^{p-1}U^*Y),$$

donde usamos que $|X^*|^{p-1} = V|X|^{p-1}V^*$, y el resultado deseado se obtiene fácilmente. □

El próximo corolario es una consecuencia inmediata del Teorema 3.2.7.

Corolario 3.2.8. *Sea $1 < p < \infty$ y $G_p : S_p \to \mathbb{R}^+$, $G_p(X) = \|X\|_p^p$. Sean $X, Y \in S_p$, $h > 0$ y $\phi \in [0, 2\pi)$, entonces*

$$D_\phi G_p(X,Y) = \lim_{h \to 0^+} \frac{G_p(X + he^{i\phi}Y) - G_p(X)}{h} = pRe\ tr(|X|^{p-1}e^{i\phi}U^*Y),$$

donde $X = U|X|$, es la descomposición polar del operador X.

La próxima proposición, permite deducir la derivada ϕ-direccional de la función $g_p(X) = \|X\|_p$, a partir del Corolario 3.2.8.

Proposición 3.2.9. *Sean $1 < p < \infty$ y $G_p : S_p \to \mathbb{R}$,*

$$G_p = \|X\|_p^p,$$

y consideremos $g_p : S_p \to \mathbb{R}$

$$g_p(X) = G_p(X)^{1/p} = \|X\|_p.$$

Entonces

$$D_\phi g_p(X,Y) = \frac{1}{p} G_p(X)^{\frac{1}{p}-1} D_\phi G_p(X,Y),$$

para cada $\phi \in [0, 2\pi)$ y para cada $X, Y \in S_p$, tal que $G_p(X) \neq 0$.

Demostración. Observar que $g_p(X) = f(G_p(X))$, con $f : \mathbb{R}_{\geq 0} \to \mathbb{R}$, $f(t) = t^{\frac{1}{p}}$. Definimos

$$v = \frac{f(t+h) - f(t)}{h} - f'(t),$$

como f es derivable para todo $t > 0$, vale que $v \underset{h \to 0}{\to} 0$, para todo $t > 0$. Y definimos

$$w = \frac{G_p(X + e^{i\phi}kY) - G_p(X)}{k} - D_\phi G_p(X, Y),$$

como existe la derivada $\phi-$ direccional de G_p para cada $\phi \in [0, 2\pi)$ y para cada $X, Y \in S_p$, (ver Corolario 3.2.8), vale que $w \underset{k \to 0^+}{\to} 0$, para todo $X, Y \in S_p$.

Usando las definiciones anteriores, tenemos que

$$f(t+h) = f(t) + [f'(t) + v]h, \tag{3.9}$$

y

$$G_p(X + e^{i\phi}kY) = G_p(X) + [D_\phi G_p(X, Y) + w]k. \tag{3.10}$$

Tomando en (3.9), $h = [D_X G_p(Y) + w]k$, (donde vale que $h = [D_X G_p(Y) + w]k \underset{k \to 0^+}{\to} 0$) y $t = G_p(X)$, para algún $X \in S_p$ tal que $G_p(X) \neq 0$, tenemos

$$f(G_p(X) + [D_\phi G_p(X, Y) + w]k) =$$

$$f(G_p(X)) + [f'(G_p(X)) + v][D_\phi G_p(X, Y) + w]k.$$

Por otra parte, usando (3.10) y la ecuación anterior, tenemos

$$f(G_p(X + e^{i\phi}kY)) = f(G_p(X) + [D_\phi G_p(X, Y) + w]k) =$$

$$= f(G_p(X)) + [f'(G_p(X)) + v][D_\phi G_p(X, Y) + w]k.$$

Entonces

$$\frac{f(G_p(X + e^{i\phi}kY)) - f(G_p(X))}{k} =$$

$$= \frac{f(G_p(X)) + [f'(G_p(X)) + v][D_\phi G_p(X, Y) + w]k - f(G_p(X))}{k} =$$

$$= [f'(G_p(X)) + v][D_\phi G_p(X, Y) + w] \underset{k \to 0^+}{\to} f'(G_p(X)) D_\phi G_p(X, Y).$$

Por lo tanto

$$D_\phi g_p(X, Y) = \frac{1}{p} G_p(X)^{\frac{1}{p} - 1} D_\phi G_p(X, Y),$$

para cada $\phi \in [0, 2\pi)$ y para cada $X, Y \in S_p$, tal que $G_p(X) \neq 0$. $\square$

Caso $p = 1$

A continuación, seguiremos [54, Teorema 2.1], para probar la existencia de la derivada ϕ- direccional de G_1 y deduciremos la fórmula de la misma.

Teorema 3.2.10. *Sea* $G_1 : S_1 \to \mathbb{R}^+$, $G_1(X) = \|X\|_1$. *Sean* $X, Y \in S_1$, *entonces*

$$lim_{h\to 0^+} \frac{\|X + hY\|_1 - \|X\|_1}{h} = Re\ tr(U^*Y) + \|QYP\|_1,$$

donde $X = U|X|$, *es la descomposición polar del operador* X, $P = P_{N(X)}$ *y* $Q = P_{N(X^*)}$.

Para la prueba de este teorema, necesitamos algunas definiciones y tres lemas técnicos.

Definición. Sea $(\mathcal{H}, \langle\cdot,\cdot\rangle)$ un espacio de Hilbert. Diremos que la sucesión $\{x_n\}_{n\geq 1} \subseteq \mathcal{H}$ converge débilmente a $x \in \mathcal{H}$ si

$$\lim_{n\to\infty} \langle\, x_n - x, y\,\rangle = 0, \text{ para cada } y \in \mathcal{H}.$$

Lema 3.2.11. *Sean* $X, Y \in L(\mathcal{H})$. *Sean* $X = U|X|$ *y* $X + hY = V_h|X + hY|$ *las descomposiciones polares de* X *y* $X + hY$ *respectivamente, sea* $Q(P)$ *el proyector al* $N(X^*)(N(X))$ *y sea* $\{\chi_j\}_{j\geq 1}$ *una base ortonormal de* $N(X)$. *Entonces*

1. $V_{h_n}x \to Ux$, *fuertemente, para todo* $x \in \overline{R(X^*)}$, *y para alguna sucesión* $h_n \to 0^+$. *También* $V^*_{h_n}x \to U^*x$, *fuertemente, para todo* $x \in \overline{R(X)}$, *y para alguna (ó la misma) sucesión* $h_n \to 0^+$,
2. $lim_{n\to\infty} \sum_j \left\langle V^*_{h_n}(I-Q)\chi_j, \chi_j \right\rangle = 0$, *si* $Y \in S_1$.

Demostración. 1. Sea $\{e_j\}_{j\geq 1}$ una base ortonormal de $\mathcal{H}$. Para cada j, la familia $\{V_h e_j : h > 0\}$ es acotada, y entonces, existe una sucesión $h_n \to 0$, tal que $V_{h_n}e_j$ converge débilmente, [19, Capítulo V, Teorema 4.2]. Más aún, usando el proceso diagonal de Cantor, concluímos que existe una sucesión $h_n \to 0$ tal que $V_{h_n}e_j$ converge débilmente para todo j, y por lo tanto V_{h_n} converge débilmente. Sea V_0 el límite débil de la sucesión V_{h_n}.

Ahora, para todo $y, z \in \mathcal{H}$ tenemos

$$\langle\, V_{h_n}|X + h_nY|z, y\,\rangle = \langle\, (X + h_nY)z, y\,\rangle\,.$$

Entonces, $X + h_nY$ converge fuertemente (incluso uniformemente) a X, lo mismo con $|X + h_nY|$ que converge fuertemente a $|X|$, y tomando límites obtenemos que

$$\langle\, V_0|X|z, y\,\rangle = \langle\, Xz, y\,\rangle = \langle\, U|X|z, y\,\rangle\,, \text{ para todo } z, y \in \mathcal{H}.$$

Entonces $V_0x = Ux$, para todo $x \in R(|X|)$. Ya que, $R(|X|)$ es denso en $\overline{R(X^*)}$, tenemos que V_{h_n} converge débilmente a U para todo $x \in \overline{R(X^*)}$. Es más esta convergencia es fuerte. De hecho, sea x un vector arbitrario de $R(X^*)$, usando la descomposición polar de X, existe $z \in \mathcal{H}$ tal que $x = |X|z$. Tenemos que $V_{h_n}|X + h_nY|z$ converge débilmente a Ux. Pero, $V_{h_n}|X + h_nY|z = (X + h_nY)z$ converge fuértemente a $Xz = Ux$. Entonces $V_{h_n}|X + h_nY|z$ converge fuértemente a Ux. Entonces tenemos que

$$\|V_{h_n}x - Ux\| \leq \|V_{h_n}(|X|z - |X + h_nY|z)\| + \|V_{h_n}|X + h_nY|z - Xz\|,$$

que tiende a 0 cuando n tiende a infinito. De manera similar, podemos obtener que $V_{h_n}^*x$ tiende a U^*x para todo $x \in \overline{R(X)}$ y para alguna (o la misma) sucesión h_n.

2. Por la parte 1. tenemos que $PV_{h_n}^*(I - Q)YP$ converge fuértemente a $PU^*(I - Q)YP$, y por [46, Teorema III.6.3], $PV_{h_n}^*(I - Q)YP$ converge a $PU^*(I - Q)YP$ en S_1, ya que $Y \in S_1$. Pero, $\sum_j \left\langle V_{h_n}^*(I - Q)Y\chi_j, \chi_j \right\rangle$ es precisamente la traza del operador $PV_{h_n}^*(I - Q)YP$ y por lo tanto tiende a la traza del operador $PU^*(I - Q)YP$. Pero $PU^* = 0$ y la prueba está completa. □

Lema 3.2.12. *Sea $T \in L(\mathcal{H})$ tal que $\|T\| \leq 1$, y sea $\{\phi_j\}_{j\geq 1}$ un sistema ortonormal arbitrario de $\mathcal{H}$. Entonces, tenemos que*

$$\|X\|_1 \geq |\sum_j \langle TX\phi_j, \phi_j \rangle|.$$

Demostración. De hecho,

$$|\sum_j \langle TX\phi_j, \phi_j \rangle| \leq |tr(TX)| \leq \|T\|\|X\|_1 \leq \|X\|_1,$$

donde se usó e Lema 3.1.11 y el Lema 3.1.13. □

Lema 3.2.13. *Sea $\{\phi_j\}_{j\geq 1}$ un sistema ortonormal (no necesariamente completo) en $\mathcal{H}$. Entonces*

1. *Para cualquier vector $f \in \mathcal{H}$, y para todo $\varepsilon > 0$, existe un vector f', tal que*

$$\|f - f'\| \leq \varepsilon \quad y \quad \sum_j |\langle f', \phi_j \rangle| < \infty,$$

2. *el conjunto*

$$F = \{T \in S_1 : \sum_j \|T\phi_j\| < \infty\},$$

es denso en S_1.

Demostración. 1. Sea $f = f_1 + f_2$ con $f_1 \in gen\{\phi_j : j \geq 1\}$ y $f_2 \in gen\{\phi_j : j \geq 1\}^\perp$.

Ya que, tenemos que $\|f_1\| = \sum_j |\langle f_1, \phi_j \rangle|^2$, dado $\varepsilon > 0$, existe $n_0 \in \mathbb{N}$, tal que $\sum_{j>n_0} |\langle f_1, \phi_j \rangle|^2 < \varepsilon^2$. Definimos

$$f' = \sum_{j \leq n_0} \langle f_1, \phi_j \rangle \phi_j + f_2,$$

entonces

$$\sum_j |\langle f', \phi_j \rangle| = \sum_{j \leq n_0} |\langle f_1, \phi_j \rangle| < \infty,$$

y también

$$\|f - f'\|^2 = \| \sum_{j>n_0} \langle f_1, \phi_j \rangle \phi_j\|^2 = \sum_{j>n_0} |\langle f_1, \phi_j \rangle|^2 < \varepsilon^2.$$

2. Sea $Y \in S_1$ y sea $\varepsilon > 0$, entonces por el Teorema 3.1.12, existe $Z \in \mathcal{F}$ tal que $\|Y - Z\|_1 < \frac{\varepsilon}{2}$. Como Z es de rango finito, podemos escribirlo de la siguiente manera $Z = \sum_{k=1}^{N} \sigma_k g_k \otimes f_k$, con $0 \leq \sigma_{k+1} \leq \sigma_k$, y f_k, g_k sistemas ortonormales. Entonces

$$\|Y - Z\|_1 < \frac{\varepsilon}{2}.$$

Por la parte 1., existen vectores f_k', tales que

$$\sum_j |\langle f_k', \phi_j \rangle| < \infty, \text{ y } \|f_k - f_k'\| < \frac{\varepsilon}{2N\sigma_k}.$$

Sea $T = \sum_{k=1}^{N} \sigma_k g_k \otimes f_k'$. Entonces, por la desigualdad triangular de la norma $\| \cdot \|_1$ y la ecuación (B.1), tenemos que

$$\|T - Z\|_1 = \| \sum_{k=1}^{N} \sigma_k g_k \otimes (f_k - f_k')\|_1 \leq \sum_{k=1}^{N} \sigma_k \|f_k - f_k'\| < \frac{\varepsilon}{2},$$

y por lo tanto

$$\|T - Y\|_1 < \varepsilon.$$

Por otra parte

$$\sum_j \|T\phi_j\| \leq \sum_{k=1}^{N} \sum_j \|\sigma_k \langle \phi_j, f_k' \rangle g_k\| = \sum_{k=1}^{N} \sigma_k \sum_j |\langle \phi_j, f_k' \rangle| < \infty.$$

$\square$

Ahora, podemos probar el Teorema 3.2.10.

Demostración. (Teorema 3.2.10.) Sea $X = \sum_j s_j \phi_j \otimes \psi_j$ la expansión del operador $X \in S_1$ tal como se expresó en (A.1) y sea $\{\chi_j\}_{j \geq 1}$ una base ortonormal de $N(X)$. Entonces, teniendo en cuenta el Lema 3.2.12, tenemos:

$$\frac{1}{h}\{\|X + hY\|_1 - \|X\|_1\} = \frac{1}{h}\{\|X + hY\|_1 - \sum_j s_j\} \geq$$

$$\frac{1}{h}(\sum_j |\langle U^*(X+hY)\phi_j,\phi_j\rangle + \sum_j \langle V^*(X+hY)\chi_j,\chi_j\rangle| - \sum_j s_j),$$

donde $V : N(X) \to N(X^*)$ está dada por $QYP = V|QYP|$.

Observar que en la ecuación anterior, el sistema ortonormal considerado es el dado por $\{\phi_j : j \geq 1\} \cup \{\chi_j : j \geq 1\}$, es decir el que se produce al unir las bases consideradas de $N(X)^\perp$ y $N(X)$ respectivamente. Además como $N(U) = N(X)$, $R(U) = \overline{R(X)}$ y que $N(X)^\perp \subseteq N(V)$, $R(V) \subseteq N(X^*) = R(X)^\perp$ vale que si $z = z_1 + z_2 \in \mathcal{H}$, con $z_1 \in N(X)$ y $z_2 \in N(X)^\perp$ entonces $\|(U+V)z\|^2 = \|Uz_2 + Vz_1\|^2 = \|Uz_2\|^2 + \|Vz_1\|^2 \leq \|z_1\|^2 + \|z_2\|^2 = \|z\|^2$, por lo tanto $\|U+V\| = \|U^* + V^*\| \leq 1$.

Entonces, como $|X| = U^*X$, y $|X|\phi_j = s_j\phi_j$, que tenemos que

$$\frac{1}{h}\{\|X+hY\|_1 - \|X\|_1\} \geq \frac{1}{h}(|\sum_j \langle |X|\phi_j,\phi_j\rangle + h\sum_j \langle U^*Y\phi_j,\phi_j\rangle +$$

$$+\sum_j \langle V^*(X+hY)\chi_j,\chi_j\rangle| - \sum_j s_j) =$$

$$= \frac{1}{h}(|\sum_j s_j + h\sum_j \langle U^*Y\phi_j,\phi_j\rangle + \sum_j \langle V^*(X+hY)\chi_j,\chi_j\rangle| - \sum_j s_j).$$

Pero, ya que $V^*Q = V^*$ y $P\chi_j = \chi_j$, la última expresión queda igual a

$$\frac{1}{h}\{\|X+hY\|_1 - \|X\|_1\} \geq$$

$$= \frac{1}{h}\{|\sum_j s_j + h\sum_j \langle U^*Y\phi_j,\phi_j\rangle + h\sum_j \langle V^*QYP\chi_j,\chi_j\rangle| - \sum_j s_j\} =$$

$$= \frac{|\sum_j s_j + h(\ tr\ (U^*Y) + \|QYP\|_1)| - \sum_j s_j}{h} \underset{h\to 0^+}{\to} Re(\ tr(U^*Y) + \|QYP\|_1),$$

y por lo tanto

$$\lim_{h\to 0^+} \frac{\|X+hY\|_1 - \|X\|_1}{h} \geq Re(\ tr(U^*Y)) + \|QYP\|_1.$$

Demostraremos la desigualdad contraria, para aquellos $Y \in F = \{T \in S_1 : \sum_j \|T\phi_j\| < \infty\}$. Será suficiente, ya que obtendremos dos funcionales acotados líneales $Y \to \underset{h\to 0^+}{lim} \frac{\|X+hY\|_1 - \|X\|_1}{h}$ e $Y \to Re(\ tr(U^*Y)) + \|QYP\|_1$, que coinciden en el conjunto F que, por el Lema 3.2.13, es denso en S_1. En lo que sigue, siempre que escribamos $\lim_{h\to 0^+} V_h$, significa $\lim_{n\to\infty} V_{h_n}$, donde $\{h_n\}_{n\geq 1}$ es la sucesión del Lema 3.2.11 (no necesitamos preocuparnos por eso, ya que por la Proposición 3.2.2, siempre existe el límite que estamos considerando). Entonces tenemos

$$\frac{1}{h}\{\|X+hY\|_1 - \|X\|_1\} =$$

$$\frac{1}{h}\{\sum_j \langle |X+hY|\phi_j,\phi_j\rangle + \sum_j \langle |X+hY|\chi_j,\chi_j\rangle - \sum_j s_j\}. \quad (3.11)$$

Sea $X+hY = V_h|X+hY|$ tal como en el Lema 3.2.11, entonces

$$\frac{1}{h}\{\sum_j \langle\, |X+hY|\chi_j, \chi_j \,\rangle\} =$$

$$= \frac{1}{h}\sum_j \langle\, V_h^*(X+hQY)\chi_j, \chi_j \,\rangle + \sum_j \langle\, V_h^*(I-Q)Y\chi_j, \chi_j \,\rangle\,,$$

y también

$$\|QYP\|_1 \geq |\sum_j \langle\, V_h^* QYP\chi_j, \chi_j \,\rangle\,| =$$

$$= |\sum_j \langle\, V_h^* QY\chi_j, \chi_j \,\rangle\,| = |\sum_j \langle\, V_h^*(X+hQY)\chi_j, \chi_j \,\rangle\,|,$$

entonces el número real $\frac{1}{h}\sum_j \langle\, |X+hY|\chi_j, \chi_j \,\rangle$ es igual a la suma de números complejos, uno de los cuales su módulo es igual o menor a $\|QYP\|_1$, y otro, cuyo módulo es, para h suficientemente chico, igual o menor a ε (Lema 3.2.11). Entonces, para h lo suficientemente chico, tenemos que

$$\sum_j \langle\, |X+hY|\chi_j, \chi_j \,\rangle \leq \|QYP\|_1 + \varepsilon.$$

Por otra parte, por la desigualdad de Jensen aplicada a la medida espectral [76, Capítulo 6, ecuación 42], tenemos:

$$\sum_j \langle\, |X+hY|\phi_j, \phi_j \,\rangle \leq \sqrt{\sum_j \langle\, |X+hY|^2\phi_j, \phi_j \,\rangle} =$$

$$= \sum_j \sqrt{s_j^2 + 2\; h\; Re\,\langle\, Y\phi_j, s_j\psi_j \,\rangle + h^2\|Y\phi_j\|^2},$$

y aplicando (3.11)

$$\frac{1}{h}\{\|X+hY\|_1 - \|X\|_1\} \leq$$

$$\frac{\sum_j \sqrt{s_j^2 + 2\; h\; Re\,\langle\, Y\phi_j, s_j\psi_j \,\rangle + h^2\|Y\phi_j\|^2} - \sum_j s_j}{h} + \|QYP\|_1 + \varepsilon =$$

$$= \sum_j \frac{\sqrt{s_j^2 + 2\; h\; Re\,\langle\, Y\phi_j, s_j\psi_j \,\rangle + h^2\|Y\phi_j\|^2} - s_j}{h} + \|QYP\|_1 + \varepsilon =$$

$$= \sum_j \frac{h\; 2\; Re\,\langle\, Y\phi_j, s_j\psi_j \,\rangle + h^2\|Y\phi_j\|^2}{h(\sqrt{s_j^2 + 2\; h\; Re\,\langle\, Y\phi_j, s_j\psi_j \,\rangle + h^2\|Y\phi_j\|^2} + s_j)} + \|QYP\|_1 + \varepsilon$$

$$\underset{h\to 0^+}{\rightarrow} \sum_j Re\,\langle\, Y\phi_j, \psi_j \,\rangle + \|QYP\|_1 + \varepsilon = \sum_j Re\,\langle\, Y\phi_j, U\phi_j \,\rangle + \|QYP\|_1 + \varepsilon =$$

$$= Re\;(tr\;(U^*Y)) + \|QYP\|_1 + \varepsilon.$$

La desigualdad

$$\left|\frac{h\;2\;Re\left\langle\, Y\phi_j, s_j\psi_j \,\right\rangle + h^2\|Y\phi_j\|^2}{h(\sqrt{s_j^2 + 2\;h\;Re\left\langle\, Y\phi_j, s_j\psi_j \,\right\rangle + h^2\|Y\phi_j\|^2} + s_j)}\right| \leq |2\;Re(\left\langle\, Y\phi_j, \psi_j \,\right\rangle| + \|Y\phi_j\|,$$

nos permite tomar límite cuando $h \to 0^+$ bajo la suma. El resultado ahora se sigue, ya que ε puede ser arbitrariamente chico. $\square$

El próximo corolario es una consecuencia inmediata del Teorema 3.2.10.

Corolario 3.2.14. *Sea $G_1 : S_1 \to \mathbb{R}^+$, $G_1(X) = \|X\|_1$. Sean $X, Y \in S_1$, $h > 0$ y $\phi \in [0, 2\pi)$, entonces*

$$D_\phi G_1(X,Y) = \lim_{h \to 0^+} \frac{G_1(X + he^{i\phi}Y) - G_1(X)}{h} = Re\; tr(e^{i\phi}U^*Y) + \|QYP\|_1,$$

donde $X = U|X|$, es la descomposición polar del operador X, $P = P_{N(X)}$ y $Q = P_{N(X^*)}$.

4 PROYECCIONES W-AUTOADJUNTAS

Dado un espacio de Hilbert $\mathcal{H}$ con producto interno $\langle\ ,\ \rangle$, el propósito de este capítulo es estudiar las proyecciones que resultan autoadjuntas, con respecto a una perturbación del producto escalar. Más precisamente, dado un operador $W \in L(\mathcal{H})^+$ se define la forma sesquilineal $\langle\ ,\ \rangle_W : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ mediante

$$\langle x, y \rangle_W = \langle Wx, y \rangle, \qquad x, y \in \mathcal{H}.$$

Se desea caracterizar al conjunto de proyecciones Q que satisfacen $\langle Qx, y \rangle_W = \langle x, Qy \rangle_W$, para todo $x, y \in \mathcal{H}$, a las cuales se las denominará W-autoadjuntas.

Si el operador $W \in GL(\mathcal{H})^+$, es fácil probar que $\langle\ ,\ \rangle_W$ también es un producto interno sobre $\mathcal{H}$ e induce una norma en $\mathcal{H}$ equivalente a la del producto $\langle\ ,\ \rangle$, por lo que $\mathcal{H}$ con el producto $\langle\ ,\ \rangle_W$ es un espacio de Hilbert.

Los primeros resultados sobre el espacio de proyecciones W-autoadjuntas, se deben a G. Corach et al. quienes estudiaron la geometría del espacio de estas proyecciones y algunas aplicaciones a diversos problemas de teoría de operadores [24, 27, 28, 25, 26, 23].

4.1 COMPATIBILIDAD: DEFINICIONES Y PROPIEDADES

Sea $\mathcal{H}$ un espacio de Hilbert con producto interno $\langle\ ,\ \rangle$ y sea $\langle\ ,\ \rangle_W$ la forma sesquilineal en $\mathcal{H}$ inducida por $W \in L(\mathcal{H})^+$.

Definición. Dado $W \in L(\mathcal{H})^+$, un operador $T \in L(\mathcal{H})$ es W-*autoadjunto* si

$$\langle Tx, y \rangle_W = \langle x, Ty \rangle_W \quad \text{para todo } x, y \in \mathcal{H}.$$

Es fácil ver que T es W-autoadjunto si y sólo si $WT = T^*W$.

Es bien sabido que, en un espacio de Hilbert, una proyección autoadjunta satisface

$$Q^* = Q \quad \Leftrightarrow \quad \|Q\| \leq 1, \tag{4.1}$$

es decir que las proyecciones ortogonales están caracterizadas por ser contractivas. En lo que se sigue se darán resultados sobre la contractividad de las proyecciones W-autoadjuntas.

Definición. Sea $\mathcal{H}$ un espacio de Hilbert y $W \in L(\mathcal{H})^+$. Un proyección $Q \in \mathcal{Q}$ es W-*contractiva* si

$$\langle Qx, Qx \rangle_W \leq \langle x, x \rangle_W \quad \text{para todo } x \in \mathcal{H}. \tag{4.2}$$

No es difícil ver que $Q \in \mathcal{Q}$ es W-contractiva si y sólo si $Q^*WQ \leq W$. Para $W \in L(\mathcal{H})^+$, G. Corach y A. Maestripieri demostraron que las proyecciones W-autoadjuntas tienen las mismas propiedades contractivas que las proyecciones autoadjuntas en espacios de Hilbert [24, Lema 3.2], en el siguiente sentido.

Proposición 4.1.1. *Sean $W \in L(\mathcal{H})^+$ y $Q \in \mathcal{Q}$. Luego, Q es W-autoadjunta si y sólo si Q es W-contractiva.*

Demostración. Si $0 \leq Q^*WQ \leq W$, como $(Q^*W^{1/2})(Q^*W^{1/2})^* = Q^*WQ \leq W = W^{1/2}(W^{1/2})^*$, por el Teorema de Douglas 2.1.1, existe una solución reducida $E \in L(\mathcal{H})$ de la ecuación $W^{1/2}X = Q^*W^{1/2}$ y cumple que $\|E\| \leq 1$. Además, por el Corolario 2.1.3, se tiene que $E \in \mathcal{Q}$ y la ecuación (4.1) asegura que $E \in \mathcal{P}$. Luego, $Q^*W = W^{1/2}EW^{1/2}$ es un un operador autoadjunto, i.e. $Q^*W = WQ$.

Recíprocamente, sea Q W-autoadjunta, notar que

$$WQ = (WQ)Q = (Q^*W)Q = Q^*WQ \in L(\mathcal{H})^+,$$

y sea $E = I - Q$, entonces también $WE = E^*WE \in L(\mathcal{H})^+$. Por lo tanto,

$$W = W(Q + E) = Q^*WQ + E^*WE \geq Q^*WQ.$$

$\square$

Definición. Sea $\mathcal{S}$ un subespacio cerrado de $\mathcal{H}$ y $W \in L(\mathcal{H})^+$. El par $(W, \mathcal{S})$ es *compatible* si existe una proyección W-autoadjunta con rango $\mathcal{S}$, i.e., si el conjunto

$$\mathcal{P}(W, \mathcal{S}) := \{Q \in \mathcal{Q} \ : \ R(Q) = \mathcal{S}, \ WQ = Q^*W\} \neq \varnothing$$

Si $\mathcal{S}$ es un subespacio de $\mathcal{H}$ y $W \in L(\mathcal{H})^+$, el subespacio W-ortogonal a $\mathcal{S}$ está dado por

$$\mathcal{S}^{\perp_W} := \{x \in \mathcal{H} \ : \ \langle Wx, s \rangle = 0 \ \text{ para todo } s \in \mathcal{S}\}.$$

Como W es autoadjunto, es fácil ver que $\mathcal{S}^{\perp_W} = W^{-1}(\mathcal{S}^\perp) = W(\mathcal{S})^\perp$.

El siguiente resultado ([24, Lema 3.2]), establece qué condiciones debe satisfacer toda proyección del conjunto $\mathcal{P}(W, \mathcal{S})$.

Lema 4.1.2 (Krein). *Sea $Q \in L(\mathcal{H})$ una proyección con $R(Q) = \mathcal{S}$. Entonces, $Q \in \mathcal{P}(W, \mathcal{S})$ si y sólo si $N(Q) \subseteq W^{-1}(\mathcal{S}^\perp)$.*

Demostración. Si Q es W-autoadjunta y $x \in N(Q)$ entonces $Q^*Wx = WQx = 0$, es decir, $Wx \in N(Q^*) = R(Q)^\perp = \mathcal{S}^\perp$. Por lo tanto, $x \in W^{-1}(\mathcal{S}^\perp)$.

Recíprocamente, si $N(Q) \subseteq W^{-1}(\mathcal{S}^\perp)$ y $x \in \mathcal{H}$,

$$Q^*Wx = Q^*WQx + Q^*W(I - Q)x = Q^*WQx$$

porque $W(I-Q)x \in \mathcal{S}^{\perp} = R(Q)^{\perp} = N(Q^*)$. Entonces, $Q^*W = Q^*WQ$ y tomando el adjunto, se tiene $WQ = Q^*WQ = Q^*W$. □

El siguiente teorema probado en [24, Proposición 3.3], nos permite caracterizar la compatibilidad del par $(W, \mathcal{S})$.

Teorema 4.1.3. *Dados $W \in L(\mathcal{H})^+$ y un subespacio cerrado $\mathcal{S}$, el par $(W, \mathcal{S})$ resulta compatible si y sólo si*

$$\mathcal{H} = \mathcal{S} + W^{-1}(\mathcal{S}^{\perp}). \tag{4.3}$$

Demostración. Si el par $(W, \mathcal{S})$ es compatible, entonces existe $Q \in \mathcal{P}(W, \mathcal{S})$, y por el Lema 4.1.2, tenemos que $N(Q) \subseteq W^{-1}(\mathcal{S}^{\perp})$, entonces

$$\mathcal{H} = R(Q) \dot{+} N(Q) = \mathcal{S} \dot{+} N(Q) \subseteq \mathcal{S} + W^{-1}(\mathcal{S}^{\perp}).$$

Recíprocamente, lamando $\mathcal{N} = \mathcal{S} \cap W^{-1}(\mathcal{S}^{\perp})$, tenemos que

$$\mathcal{H} = \mathcal{S} + W^{-1}(\mathcal{S}^{\perp}) = \mathcal{S} \dot{+} (W^{-1}(\mathcal{S}^{\perp}) \ominus \mathcal{N}).$$

La proyección Q, definida por esta descomposición de $\mathcal{H}$ cumple que $N(Q) = W^{-1}(\mathcal{S}^{\perp}) \ominus \mathcal{N} \subseteq W^{-1}(\mathcal{S}^{\perp})$, entonces nuevamente por el Lema 4.1.2, tenemos que el par $(W, \mathcal{S})$ es compatible. □

A continuación se presentan algunos ejemplos de pares compatibles y no compatibles.

1. Si $W \in GL(\mathcal{H})^+$ entonces $(W, \mathcal{S})$ es compatible. Esto resulta evidente ya que, como se mencionó anteriormente, $\mathcal{H}$ con el producto $\langle \ , \ \rangle_W$ es un espacio de Hilbert. Más aún, en este caso existe una única proyección en el conjunto $\mathcal{P}(W, \mathcal{S})$.

2. Menos trivial resulta el siguiente resultado (ver [24, Teorema 6.2]). Si $W \in L(\mathcal{H})^+$ y $\mathcal{S}$ es un subespacio de dimensión finita, entonces $\mathcal{P}(W, \mathcal{S}) \neq \emptyset$.

El siguiente ejemplo de un par $(W, \mathcal{S})$ no compatible, en donde el operador $W \in L(\mathcal{H})^+$ es un operador con buenas propiedades, en particular es fácil ver que es de traza.

Ejemplo 4.1.4. Sea $\{e_k\}_{k=0,1,2,\ldots}$ una base ortonormal de $\mathcal{H}$, $a_k = \sqrt{k}3^{-k}$, $w_k = k3^{-k}$, $k = 1, 2, \ldots$ y $\mathcal{S}^{\perp}$ el subespacio generado por e_0, entonces $\mathcal{S} = gen\{\{e_k\}_{k=1,2,\ldots}\}$. Se define el operador $W \in L(\mathcal{H})^+$ de la siguiente manera,

$$\langle We_k, e_l \rangle = \begin{bmatrix} 1 & a_1 & a_2 & a_3 & \ldots \\ \overline{a_1} & w_1 & 0 & 0 & \ldots \\ \overline{a_2} & 0 & w_2 & 0 & \ldots \\ \vdots & & & & \ddots \end{bmatrix}, \quad k, l = 0, 1, 2, \ldots \tag{4.4}$$

no es difícil ver que W es un operador positivo y de traza.
Sea $x \in W(\mathcal{S})^{\perp}$, entonces para todo $k = 1, 2, 3, ...$

$$0 = \langle We_k, x \rangle = \left\langle We_k, \sum_{l=0} \langle x, e_l \rangle e_l \right\rangle = \sum_{l=0} \overline{\langle x, e_l \rangle} \langle We_k, e_l \rangle .$$

Entonces, de la ecuación (4.4) se sigue que, para $k = 1, 2, 3, ...,$

$$\overline{\langle x, e_0 \rangle} a_k + \overline{\langle x, e_k \rangle} w_k = 0,$$

ó equivalentemente,

$$\langle x, e_k \rangle = - \langle x, e_0 \rangle \frac{a_k}{w_k}, \text{ para todo } k = 1, 2, 3, ...$$

Entonces,

$$\|x\|^2 = \sum_{k=0} | \langle x, e_k \rangle |^2 = | \langle x, e_0 \rangle |^2 + \sum_{k=1} | \langle x, e_0 \rangle \frac{a_k}{w_k} |^2 =$$

$$= | \langle x, e_0 \rangle |^2 (1 + \sum_{k=1} | \frac{a_k}{w_k} |^2)$$

Por lo que, x tiene norma finita si y sólo si $\langle x, e_0 \rangle = 0$, es decir, $x = 0$. Entonces $\mathcal{S} + W(\mathcal{S})^{\perp} = \mathcal{S} \neq \mathcal{H}$, es decir, el par $(W, \mathcal{S})$ no es compatible.

Dados un subespacio cerrado $\mathcal{S}$ de $\mathcal{H}$ y un operador $W \in L(\mathcal{H})^{+}$, la compatibilidad del par $(W, \mathcal{S})$ puede relacionarse con los ángulos de ciertos subespacios.

Teorema 4.1.5. *([27, Teorema 2.15]) Sean $W \in L(\mathcal{H})^{+}$ y $\mathcal{S}$ un subespacio cerrado de $\mathcal{H}$. Entonces, el par $(W, \mathcal{S})$ es compatible si y sólo si $c_0(\mathcal{S}^{\perp}, \overline{W(\mathcal{S})}) < 1$.*

Demostración. Si el par $(W, \mathcal{S})$ es compatible, se tiene que $\mathcal{H} = \mathcal{S} + W(\mathcal{S})^{\perp}$, tomando el complemento ortogonal, se sigue que $\mathcal{S}^{\perp} \cap \overline{W(\mathcal{S})} = \{0\}$ entonces, $c_0(\mathcal{S}, \overline{W(\mathcal{S})}) = c(\mathcal{S}, \overline{W(\mathcal{S})})$. Por la Proposición 2.1.4, la igualdad $\mathcal{H} = \mathcal{S} + W(\mathcal{S})^{\perp}$ implica que el subespacio $\mathcal{S}^{\perp} + \overline{W(\mathcal{S})}$ es cerrado, o equivalentemente, $c(\mathcal{S}^{\perp}, \overline{W(\mathcal{S})}) < 1$. Por lo tanto, $c_0(\mathcal{S}^{\perp}, \overline{W(\mathcal{S})}) < 1$.

Recíprocamente, si $c_0(\mathcal{S}^{\perp}, \overline{W(\mathcal{S})}) < 1$, como $c(\mathcal{S}^{\perp}, \overline{W(\mathcal{S})}) \leq c_0(\mathcal{S}^{\perp}, \overline{W(\mathcal{S})})$, se tiene que $c(\mathcal{S}^{\perp}, \overline{W(\mathcal{S})}) < 1$ y además por la definición de c_0, tenemos que $\mathcal{S}^{\perp} \cap \overline{W(\mathcal{S})} = \{0\}$. Entonces, utilizando la Proposición 2.1.4, se sigue que $\mathcal{S} + W(\mathcal{S})^{\perp}$ es cerrado, y además,

$$\mathcal{S} + W(\mathcal{S})^{\perp} = \overline{\mathcal{S} + W(\mathcal{S})^{\perp}} = (\mathcal{S}^{\perp} \cap \overline{W(\mathcal{S})})^{\perp} = \{0\}^{\perp} = \mathcal{H},$$

es decir, el par $(W, \mathcal{S})$ es compatible. $\square$

Cuasi-compatibilidad

En este sección, estudiamos proyecciones no acotadas que son simétricas para el semi-producto interno definido por el operador $W \in L(\mathcal{H})^+$. También, caracterizamos la existencia de dichas proyecciones con un rango $\mathcal{S} \subseteq \mathcal{H}$ determinado. Para un resumen de las propiedades y definiciones más relevantes de operadores no acotados en un espacio de Hilbert, ver la sección 2.1.3.

Definición. Sea Q un operador densamente definido. Diremos que Q es una proyección densamente definida, si Q es idempotente, es decir

$$R(Q) \subseteq D(Q), \text{ y } Q^2x = Qx, \text{ para cada } x \in D(Q).$$

En este caso

$$D(Q) = R(Q) \dot{+} N(Q).$$

Además la proyección Q resulta cerrada si y sólo si $R(Q)$ y $N(Q)$ son subespacios cerrados.

Definición. Sea $W \in L(\mathcal{H})^+$ y Q una proyección cerrada densamente definida; diremos que Q es W-simétrica si WQ es simétrica y diremos que es W- autoadjunta si WQ es autoadjunta.

Dado que W es acotado, entonces $D(WQ) = D(Q)$ y $(WQ)^* = Q^*W$, ver Teorema 2.1.7. Por lo tanto Q es W-simétrica, si y sólo si

$$WQx = Q^*Wx \text{ para cada } x \in D(Q),$$

y Q es W-autoadjunta si y sólo si $WQ = Q^*W$.

Proposición 4.1.6. *Sean $W \in L(\mathcal{H})^\perp$, $\mathcal{S} \subseteq \mathcal{H}$ un subespacio cerrado y Q una proyección densamente definida con $R(Q) = \mathcal{S}$. Entonces Q es W-simétrica si y sólo si $N(Q) \subseteq (W\mathcal{S})^\perp$.*

Demostración. Ver [21, Proposición 2.2]. □

Dados $W \in L(\mathcal{H})^+$ y un subespacio cerrado $\mathcal{S} \subseteq \mathcal{H}$, como vimos, la teoría de compatibilidad estudia la existencia de proyecciones acotadas W-autoadjuntas con rango $\mathcal{S}$. Más precisamente, el par $(W, \mathcal{S})$ se dice compatible si existe $Q \in L(\mathcal{H})$ con rango $\mathcal{S}$ tal que $WQ = Q^*W$. Este hecho, es equivalente a $\mathcal{H} = \mathcal{S} + (W\mathcal{S})^\perp$, como se puede ver en el Teorema 4.1.3. A continuación, caracterizaremos la compatibilidad del par $(W, \mathcal{S})$ en términos de proyecciones no acotadas.

Definición. Sean $W \in L(\mathcal{H})^+$ y $\mathcal{S} \subseteq \mathcal{H}$, un subespacio cerrado. Diremos que el par $(W, \mathcal{S})$ es *cuasi-compatible* si existe una proyección W-simétrica Q tal que $R(Q) = \mathcal{S}$.

Proposición 4.1.7. *Dados $W \in L(\mathcal{H})^+$ y $\mathcal{S} \subseteq \mathcal{H}$, un subespacio cerrado. El par $(W, \mathcal{S})$ es cuasi-compatible si y sólo si*

$$\overline{\mathcal{S} + (W\mathcal{S})^{\perp}} = \mathcal{H}.$$

Demostración. Si el par $(W, \mathcal{S})$ es cuasi-compatible entonces existe una proyección W-simétrica densamente definida Q con rango $\mathcal{S}$, y por la Propoposición 4.1.6, $D(Q) = \mathcal{S} \dot{+} N(Q) \subseteq \mathcal{S} + (W\mathcal{S})^{\perp}$. Como Q es densamente definida, entonces tenemos que

$$\overline{\mathcal{S} + (W\mathcal{S})^{\perp}} = \overline{D(Q)} = \mathcal{H}.$$

Recíprocamente, sea $\mathcal{N} = \mathcal{S} \cap (W\mathcal{S})^{\perp}$. Notar que

$$\mathcal{S} + (W\mathcal{S})^{\perp} = \mathcal{S} \dot{+} (W\mathcal{S})^{\perp} \ominus \mathcal{N},$$

y definamos

$$Q = P_{\mathcal{S}//(W\mathcal{S})^{\perp} \ominus \mathcal{N}}.$$

Entonces Q es una proyección densamente definida y como $N(Q) = (W\mathcal{S})^{\perp} \ominus \mathcal{N} \subseteq (W\mathcal{S})^{\perp}$, por la Proposición 4.1.6, Q es W-simétrica. Por lo tanto, el par $(W, \mathcal{S})$ es cuasi-compatible. □

4.2 COMPATIBILIDAD: APLICACIONES

4.2.1 *Problema de cuadrados mínimos con peso*

A continuación daremos la definición de la solución de cuadrados mínimos con peso W de la ecuación $Az = x$.

Definición. Dados $A \in CR(\mathcal{H})$, $W \in L(\mathcal{H})^+$ y $x \in \mathcal{H}$, $u \in \mathcal{H}$ es una solución de cuadrados mínimos con peso W ó W-LSS de $Az = x$, si

$$\|Au - x\|_W \leq \|Az - x\|_W, \text{ para cada } z \in \mathcal{H}. \tag{4.5}$$

El próximo teorema describe algunas propiedades de la solución de cuadrados mínimos con peso W de la ecuación $Az = x$ (ver [22]).

Teorema 4.2.1. *Dados $A \in CR(\mathcal{H})$, $W \in L(\mathcal{H})^+$ y $x \in \mathcal{H}$, entonces las siguientes condiciones se siguen:*

1. *Existe una W-LSS de $Az = x$ si y sólo si $x \in R(A) + R(A)^{\perp_W}$,*
2. *u_0 es una W-LSS de $Az = x$ si y sólo si*

$$A^*W(Au_0 - x) = 0.$$

Demostración. 1. Sea u_0 una W-LSS de $Az = x$. Sea $v = Au_0 - x$, veamos que $v \in R(A)^{\perp_W}$. Sea $z \in \mathcal{H}$, entonces, para cualquier $t \in \mathbb{R}$

$$\|v\|_W^2 \leq \|A(tz + u_0) - x\|_W^2 = \|Atz + v\|_W^2 = \langle W(Atz + v), Atz + v \rangle = \\ \|v\|_W^2 + t^2 \langle WAz, Az \rangle + 2t \, Re \langle Wv, Az \rangle ,$$

entonces,

$$t^2 \langle WAz, Az \rangle + 2t \, Re \langle Wv, Az \rangle \geq 0, \text{ para todo } t \in \mathbb{R}.$$

Usando un argumento estándar obtenemos que $\langle Wv, Az \rangle = 0$, entonces $v \in R(A)^{\perp_W}$. Por lo tanto $x = Au_0 - v \in R(A) + R(A)^{\perp_W}$.
Recíprocamente si $x \in R(A) + R(A)^{\perp_W}$, existen $u_0 \in \mathcal{H}$ y $v \in R(A)^{\perp_W}$ tal que $x = Au_0 + v$. Entonces para todo $h \in \mathcal{H}$, vale $\langle Wv, Ah \rangle = \langle WAh, v \rangle = 0$. Sea $z \in \mathcal{H}$, entonces

$$\|x - Az\|_W^2 = \|v + A(u_0 - z)\|_W^2 = \langle W(v + A(u_0 - z)), v + A(u_0 - z) \rangle = \\ = \|v\|_W^2 + \|A(u_0 - z)\|_W^2 \geq \|v\|_W^2 = \|x - Au_0\|_W^2.$$

Por lo tanto u_0 es una W-LSS de $Az = x$.
2. Dado $x \in \mathcal{H}$, u satisface (4.5) si y sólo si

$$\|W^{1/2}(x - Au)\| \leq \|W^{1/2}(x - Az)\| =$$

$$= \|W^{1/2}(x - Au) + W^{1/2}A(u - z)\|, \text{ para todo } z \in \mathcal{H},$$

ó equivalentemente

$$\left\langle W^{1/2}x - W^{1/2}Au, W^{1/2}Az \right\rangle = 0, \text{ para todo } z \in \mathcal{H}.$$

Es fácil ver que dados $x, y \in \mathcal{H}$, tenemos que

$$\|x\| \leq \|x + ty\| \text{ para todo } t \in \mathbb{C},$$

si y sólo si

$$\langle x, y \rangle = 0.$$

Entonces u es una W-LSS de $Az = x$ si y sólo si u es solución de la ecuación normal

$$A^*WAy = A^*Wx. \tag{4.6}$$

□

El próximo teorema, relaciona la existencia de W-LSS de la ecuación $Az = x$ con la compatibilidad del par $(W, R(A))$.

Teorema 4.2.2. *Sea $A \in CR(\mathcal{H})$, $W \in L(\mathcal{H})^+$, entonces existe u una W-LSS de la ecuación $Az = x$ para cada $x \in \mathcal{H}$ si y sólo si el par $(W, R(A))$ es compatible.*

Demostración. Supongamos que existe u una W-LSS de la ecuación $Az = x$, para cada $x \in \mathcal{H}$. Entonces si $x \in \mathcal{H}$, por el Teorema 4.2.1, $x \in R(A) + R(A)^{\perp_W}$, es decir $\mathcal{H} = R(A) + R(A)^{\perp_W}$ y por el Teorema 4.1.3, el par $(W, R(A))$ es compatible.

Recíprocamente, si el par $(W, R(A))$ es compatible, por el Teorema 4.1.3, $\mathcal{H} = R(A) + R(A)^{\perp_W}$. Entonces, si $x \in \mathcal{H}$, $x \in R(A) + R(A)^{\perp_W}$, y por el Teorema 4.2.1, existe u una W-LSS de la ecuación $Az = x$, para cada $x \in \mathcal{H}$. □

4.2.2 *W-inversas*

En [72] S. K. Mitra y C. R. Rao introdujeron la noción de W-inversa de una matriz. Ahora extenderemos la definición en el siguiente sentido.

Definición. Dados $A \in CR(\mathcal{H})$, $B \in L(\mathcal{H})$ y $W \in L(\mathcal{H})^+$, $X_0 \in L(\mathcal{H})$ es una W-inversa de A en $R(B)$, si para cada $x \in \mathcal{H}$, $X_0 x$ es una W-LSS de $Az = Bx$, es decir

$$\|AX_0 x - Bx\|_W \leq \|Az - Bx\|_W, \text{ para cada } z \in \mathcal{H}.$$

Cuando $B = I$, X_0 se llama la W-inversa de A (ver [20]). El próximo teorema muestra que existe una estrecha relación entre las W-inversas y las soluciones W-LSS.

Teorema 4.2.3. *Dados $A \in CR(\mathcal{H}), B \in L(\mathcal{H})$ y $W \in L(\mathcal{H})^+$, las siguientes condiciones son equivalentes:*

1. *El operador A admite una W-inversa en $R(B)$,*
2. $R(B) \subseteq R(A) + R(A)^{\perp_W}$,
3. *la ecuación normal $A^*WAX = A^*WB$ admite una solución.*

Demostración. *1.* $\Leftrightarrow$ *3.:* Si X_0 es una W-inversa de A en $R(B)$ entonces,

$$\|AX_0 x - Bx\|_W \leq \|Az - Bx\|_W, \text{ para cada } x, z \in L(\mathcal{H}),$$

ó equivalentemente, $X_0 x$ es una W-LSS de $Az = Bx$, para cada $x \in \mathcal{H}$, ó, por el Teorema 4.2.1,

$$A^*W(AX_0 - B)x = 0, \text{ para cada } x \in \mathcal{H},$$

entonces X_0 es una solución de la ecuación normal. La recíproca se sigue de manera similar, aplicando el Teorema 4.2.1.
2. $\Leftrightarrow$ *3.:* Si $R(B) \subseteq R(A) + R(A)^{\perp_W}$, aplicando A^*W a ambos lados de la inclusión obtenemos,

$$R(A^*WB) \subseteq R(A^*WA).$$

Luego por el Teorema 2.1.1, la ecuación normal admite solución.

Recíprocamente, si la ecuación normal $A^*WAX = A^*WB$ admite una solución, entonces tenemos que

$$A^*WR(B) \subseteq A^*WR(A),$$

aplicando $(A^*W)^{-1}$ a ambos lados de la inclusión obtenemos

$$R(B) \subseteq (A^*W)^{-1}(A^*WR(B)) \subseteq$$

$$(A^*W)^{-1}(A^*WR(A)) = R(A) + N(A^*W) = R(A) + W^{-1}N(A^*) =$$
$$= R(A) + W^{-1}(R(A)^{\perp}) = R(A) + R(A)^{\perp_W}.$$

$\square$

Observar en el teorema anterior, que si $B = I$, entonces el operador A admite una W-inversa si y sólo si el par $(W, R(A))$ es compatible.

Corolario 4.2.4. *Si $R(B) \subseteq R(A) + R(A)^{\perp_W}$, entonces el conjunto de W-inversas de A en $R(B)$ es el conjunto de soluciones de la ecuación normal $A^*WAX = A^*WB$, ó equivalentemente la variedad afín*

$$(A^*WA)^{\dagger}A^*WB + \{L \in L(\mathcal{H}) : R(L) \subseteq N(A^*WA)\}.$$

5 OPERADOR DE SHORTED Y COMPRESIONES

Dado un subespacio cerrado $\mathcal{S} \subseteq \mathcal{H}$, en esta sección usaremos el hecho de que $P = P_{\mathcal{S}} \in \mathcal{P}$ induce una representación matricial de todo operador $T \in L(\mathcal{H})$, en matrices de 2×2. Si $T \in L(\mathcal{H})$ se descompone como

$$T = PTP + PT(I-P) + (I-P)TP + (I-P)T(I-P),$$

entonces T está representado por la matriz

$$T = \begin{bmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{bmatrix}, \tag{5.1}$$

donde $t_{11} : \mathcal{S} \to \mathcal{S},\ t_{21} : \mathcal{S} \to \mathcal{S}^{\perp},\ t_{12} : \mathcal{S}^{\perp} \to \mathcal{S},\ t_{22} : \mathcal{S}^{\perp} \to \mathcal{S}^{\perp}$.

Bajo esta representación P puede ser identificada como

$$P = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix},$$

y todas las matrices idempotentes $Q \in \mathcal{Q}$ con el mismo rango que P tienen la forma

$$Q = \begin{bmatrix} I & x \\ 0 & 0 \end{bmatrix},$$

para algún $x \in L(N(P), R(P))$.

5.1 OPERADOR DE SHORTED

Sea $W \in L(\mathcal{H})^+$ y $\mathcal{S} \subseteq \mathcal{H}$ un subespacio cerrado, la noción de operador de *shorted* de W a $\mathcal{S}$, fue introducida por M.G. Krein [55] y fue luego redescubierta por W.N.Anderson y G.E. Trapp [3], [4], quienes lo aplicaron a la teoría de redes eléctricas.

En espacios de dimensión finita, el operador de shorted es una de las varias manifestaciones del complemento de Schur de una matriz. Dada una matrix en bloque

$$W = \begin{bmatrix} W_{11} & W_{12} \\ W_{12}^* & W_{22} \end{bmatrix},$$

con W_{11} inversible. $W_{22} - W_{12}^* W_{11}^{-1} W_{12}$ es el complemento de Schur de W_{11} en W. Esta definición se debe a E. Haynsworth [52].

En [10] y [30] se recopilan varias propiedades y aplicaciones. La noción fue generalizada en varias direcciones. En particular, T. Ando [5] introdujo, simultáneamente junto con una generalización del complemento de Schur, el concepto de $\mathcal{S}$-compresión. En la definición de Ando, si $\mathcal{S}$ es un subespacio de $\mathcal{H}$ y W es un operador positivo en $\mathcal{H}$ de la forma

$$W = \begin{bmatrix} W_{11} & W_{12} \\ W_{12}^* & W_{22} \end{bmatrix},$$

con W_{11} inversible en $\mathcal{S}$, entonces $W_{/\mathcal{S}} = \begin{bmatrix} 0 & 0 \\ 0 & W_{22} - W_{12}^* W_{11}^{-1} W_{12} \end{bmatrix}$, y $W_{\mathcal{S}} = \begin{bmatrix} W_{11} & W_{12} \\ W_{12}^* & W_{12}^* W_{11}^{-1} W_{12} \end{bmatrix}$.

W.N. Anderson y G. E. Trapp en [3], caracterizaron el complemento de Schur como el máximo del conjunto

$$\{X \in L(\mathcal{H}) : 0 \leq X \leq W \text{ y } R(X) \subseteq \mathcal{S}^\perp\}.$$

Esta propiedad se usó por dichos autores, para extender el concepto de operador de shorted a operadores positivos en espacios de Hilbert arbitrarios, [4].

Teorema 5.1.1 ([4]). *Sea $W \in L(\mathcal{H})^+$ y $\mathcal{S} \subseteq \mathcal{H}$ un subespacio cerrado. Sea el conjunto $\mathcal{M}(W, \mathcal{S})$ de operadores positivos definido por*

$$\mathcal{M}(W, \mathcal{S}) = \{X \in L(\mathcal{H}) : 0 \leq X \leq W \text{ y } R(X) \subseteq \mathcal{S}^\perp\}.$$

Entonces $\mathcal{M}(W, \mathcal{S})$ tiene un elemento máximo.

Demostración. Primero observar que debido a que $A \leq B$ y $B \leq A$ implica $A = B$, si $\mathcal{M}(W, \mathcal{S})$ tiene un elemento máximo, este máximo es único.

Ahora, supongamos que W es inversible. Entonces la matriz de W inducida por $\mathcal{S}$ es

$$W = \begin{bmatrix} W_{11} & W_{12} \\ W_{12}^* & W_{22} \end{bmatrix}, \tag{5.2}$$

donde $W_{11} : \mathcal{S} \to \mathcal{S}$, $W_{12}^* : \mathcal{S} \to \mathcal{S}^\perp$, $W_{12} : \mathcal{S}^\perp \to \mathcal{S}$, $W_{22} : \mathcal{S}^\perp \to \mathcal{S}^\perp$. Como W es inversible y positivo, también lo es W_{11} y el operador shorted $W_{/\mathcal{S}}$ está dado por la matriz en bloques

$$W_{/\mathcal{S}} = \begin{bmatrix} 0 & 0 \\ 0 & W_{22} - W_{12}^* W_{11}^{-1} W_{12} \end{bmatrix}. \tag{5.3}$$

Dado un vector arbitrario en $\mathcal{H}$, este puede ser escrito en la forma $\begin{bmatrix} x \\ y \end{bmatrix}$, con $x \in \mathcal{S}$ e $y \in \mathcal{S}^\perp$. Sea $z = x + W_{11}^{-1} W_{12} y$. Luego

$$\langle W \begin{bmatrix} x \\ y \end{bmatrix}, \begin{bmatrix} x \\ y \end{bmatrix} \rangle = \langle W_{11} x, x \rangle + \langle W_{12} y, x \rangle + \langle W_{12}^* x, y \rangle + \langle W_{22} y, y \rangle =$$

$$= \langle W_{11}z, z\rangle + \langle W_{22} - W_{12}^* W_{11}^{-1} W_{12} y, y\rangle \geq \langle W_{/\mathcal{S}} \begin{bmatrix} x \\ y \end{bmatrix}, \begin{bmatrix} x \\ y \end{bmatrix}\rangle,$$

donde se usó que como W es positivo, también lo es W_{11}; por lo tanto $W_{/\mathcal{S}} \leq W$. Más aún, para cualquier $y \in \mathcal{H}$ podemos tomar x tal que $z = 0$; por lo tanto $W_{/\mathcal{S}} \geq 0$. Entonces $W_{/\mathcal{S}} \in \mathcal{M}(W, \mathcal{S})$.

A continuación, supongamos que D es un operador positivo tal que $D \leq W$ y $R(D) \subseteq \mathcal{S}^\perp$. Luego, respecto de la misma partición usada arriba tenemos

$$D = \begin{bmatrix} 0 & 0 \\ 0 & D_{22} \end{bmatrix},$$

y para cualquier vector en $\mathcal{H}$ tenemos

$$\langle D \begin{bmatrix} x \\ y \end{bmatrix}, \begin{bmatrix} x \\ y \end{bmatrix}\rangle = \langle D \begin{bmatrix} -W_{11}^{-1} W_{12} y \\ y \end{bmatrix}, \begin{bmatrix} -W_{11}^{-1} W_{12} y \\ y \end{bmatrix}\rangle \leq$$

$$\langle W \begin{bmatrix} -W_{11}^{-1} W_{12} y \\ y \end{bmatrix}, \begin{bmatrix} -W_{11}^{-1} W_{12} y \\ y \end{bmatrix}\rangle = \langle W_{/\mathcal{S}} \begin{bmatrix} x \\ y \end{bmatrix}, \begin{bmatrix} x \\ y \end{bmatrix}\rangle.$$

Por lo tanto $W_{/\mathcal{S}}$ es el máximo elemento de $\mathcal{M}(W, \mathcal{S})$, y el teorema queda probado para W inversible.

Para construir el máximo en el caso en que W no sea inversible, usaremos una técnica de límites. Observar que si $Y \geq W$ y W es inversible, entonces $Y_{/\mathcal{S}} \geq W_{/\mathcal{S}}$, ya que $Y_{/\mathcal{S}}$ es el máximo sobre un conjunto más grande. Sea $\{E_n\}_{n\geq 1}$ una sucesión monótona decreciente (en el orden de operadores) de operadores positivos que convergen fuertemente a 0. Entonces $W + E_n$ es inversible para todo n, y para $n \geq m$, vale que $W + E_n \leq W + E_m$. Entonces la sucesión $\{(W + E_n)_{/\mathcal{S}}\}_{n\geq 1}$ es monótona decreciente y por lo tanto tiene un límite fuerte F, [74, Teorema pág. 263]. Veamos que F es el máximo de $\mathcal{M}(W, \mathcal{S})$. Primero observemos que para todo n,

$$F \leq (W + E_n)_{/\mathcal{S}} \leq W + E_n;$$

y tomando límite tenemos que $F \leq W$. Más aún, para todo n

$$R((W + E_n)_{/\mathcal{S}}) \subseteq \mathcal{S}^\perp,$$

entonces el límite F también cumple $R(F) \subseteq \mathcal{S}^\perp$. Por lo tanto $F \in \mathcal{M}(W, \mathcal{S})$.

Segundo, supongamos que $D \in \mathcal{M}(W, \mathcal{S})$. Entonces $D \leq W$ y $D \leq W + E_n$, por lo tanto $D \leq (W + E_n)_{/\mathcal{S}}$, para todo n. Tomando límite, obtenemos $D \leq F$. Entonces, F es el máximo de $\mathcal{M}(W, \mathcal{S})$. □

Definición. Sea $W \in L(\mathcal{H})^+$ y $\mathcal{S} \subseteq \mathcal{H}$ un subespacio cerrado. El operador de shorted $W_{/\mathcal{S}}$, es el máximo garantizado por el teorema anterior. Observar que debido a que $A \leq B$ y $B \leq A$ implica $A = B$, este máximo es único.

La $\mathcal{S}$-compresión $W_{\mathcal{S}}$ se define como $W_{\mathcal{S}} = W - W_{/\mathcal{S}}$.

En lo que sigue, $\mathcal{S} \subseteq \mathcal{H}$ es un subespacio cerrado y usaremos para W la descomposición (5.1).

Corolario 5.1.2. *Sean $W, Y \in L(\mathcal{H})^+$ tal que $W \leq Y$. Entonces $W_{/\mathcal{S}} \leq Y_{/\mathcal{S}}$.*

Demostración. $\mathcal{M}(W, \mathcal{S}) \subseteq \mathcal{M}(Y, \mathcal{S})$ y como $W_{/\mathcal{S}}$ e $Y_{/\mathcal{S}}$ son los máximos correspondientes, tenemos que $W_{/\mathcal{S}} \leq Y_{/\mathcal{S}}$. □

Corolario 5.1.3. *Sea $\{E_n\}_{n \geq 1}$ una sucesión monótona decreciente de operadores convergentes a 0 en sentido fuerte, y sea $W \in L(\mathcal{H})^+$; entonces $(W + E_n)_{/\mathcal{S}}$ converge en sentido fuerte a $W_{/\mathcal{S}}$.*

Demostración. Sea $F_n = E_n + \frac{1}{n}$. Entonces $W \leq W + E_n \leq W + F_n$. Ahora $W + F_n$ es inversible, entonces podemos aplicar la prueba del teorema anterior, y obtenemos

$$W_{/\mathcal{S}} \leq (W + E_n)_{/\mathcal{S}} \leq (W + F_n)_{/\mathcal{S}}.$$

El resultado se sigue del hecho que $(W + F_n)_{/\mathcal{S}}$ converge en sentido fuerte a $W_{/\mathcal{S}}$. □

El próximo conocido teorema, caracteriza los operadores positivos a partir de la descomposición matricial presentada en (5.1). La prueba de este resultado está basada en la existencia del operador de shorted [4, Teorema 3].

Teorema 5.1.4. *Sea $\mathcal{S} \subseteq \mathcal{H}$ un subespacio cerrado y $T \in L(\mathcal{H})$ con la descomposición matricial dada en (5.1). Entonces $T \in L(\mathcal{H})^+$ si y sólo si*

1. $t_{12} = t_{21}^*$,
2. $t_{11} \geq 0$,
3. $R(t_{12}) \subseteq R(t_{11}^{1/2})$,
4. $t_{22} = ((t_{11}^{1/2})^\dagger t_{12})^*(t_{11}^{1/2})^\dagger t_{12} + f$, *con* $f \geq 0$.

Ahora enunciamos, algunos resultados de Anderson-Trapp y E.L. Pekarev [71], los cuales son relevantes para este trabajo.

Teorema 5.1.5 ([4]). *Sea $W \in L(\mathcal{H})^+$ y $\mathcal{S} \subseteq \mathcal{H}$ un subespacio cerrado. Sea $\mathcal{N}(W, \mathcal{S})$ el conjunto de operadores*

$$\mathcal{N}(W, \mathcal{S}) = \{E^*WE : E \in \mathcal{Q},\ N(E) = \mathcal{S}\},$$

entonces el conjunto $\mathcal{N}(W, \mathcal{S})$ tiene ínfimo. Más aún

$$W_{/\mathcal{S}} = inf\{E^*WE : E \in \mathcal{Q},\ N(E) = \mathcal{S}\};$$

en general, el ínfimo no se alcanza.

Demostración. Primero, supongamos que W es inversible. Entonces usando la misma representación matricial para W que en el Teorema 5.1.1, como W es inversible, también lo es W_{11}. Sea $T \in L(\mathcal{H})$ un operador arbitrario, entonces podemos escribir la matriz de cualquier $E \in \mathcal{Q}$ con $N(E) = \mathcal{S}$ como

$$E = \begin{bmatrix} 0 & T - W_{11}^{-1}W_{12} \\ 0 & I \end{bmatrix}.$$

Entonces

$$E^*WE = \begin{bmatrix} 0 & 0 \\ 0 & W_{22} - W_{12}^*W_{11}^{-1}W_{12} + T^*W_{11}T \end{bmatrix} = W_{/\mathcal{S}} + \begin{bmatrix} 0 & 0 \\ 0 & T^*W_{11}T \end{bmatrix}.$$

Como T puede ser 0, el teorema vale para W inversibles. En este caso, el ínfimo es de hecho un mínimo.

En el caso en que W no sea inversible, sea $\varepsilon > 0$, y consideremos el operador positivo $\varepsilon I \in GL(\mathcal{H})^+$. Entonces para cualquier proyección $E \in L(\mathcal{H})$ con $N(E) = \mathcal{S}$, tenemos

$$W_{/\mathcal{S}} \leq (W + \varepsilon I)_{/\mathcal{S}} \leq E^*(W + \varepsilon I)E,$$

como ε es arbitrario, $W_{/\mathcal{S}}$ es una cota inferior de $\mathcal{N}(W, \mathcal{S})$.

Si $C \geq 0$ es otra cota inferior de $\mathcal{N}(W, \mathcal{S})$, entonces para cualquier $\varepsilon > 0$ y cualquier proyección $E \in L(\mathcal{H})$ con $N(E) = \mathcal{S}$, tenemos

$$C \leq E^*WE \leq E^*(W + \varepsilon I)E.$$

Dado que $W + \varepsilon I$ es inversible, se sigue que $C \leq (W + \varepsilon I)_{/\mathcal{S}}$, y ya que ε es arbitrario, por el Corolario 5.1.3, concluimos que $C \leq W_{/\mathcal{S}}$. $\square$

Teorema 5.1.6 ([4], [71]). *Sea $W \in L(\mathcal{H})^+$ con representación matricial (respecto de $\mathcal{S}$) $W = \begin{bmatrix} W_{11} & W_{12} \\ W_{12}^* & W_{22} \end{bmatrix}$.*

1. *Por el Teorema 5.1.4, como $W \in L(\mathcal{H})^+$, tenemos que $R(W_{12}) \subseteq R(W_{11}^{1/2})$. Sea $d \in L(\mathcal{S}^\perp, \mathcal{S})$ la solución reducida de la ecuación $W_{11}^{1/2}x = W_{12}$ (garantizada por el Teorema de Douglas 2.1.1) entonces*

$$W_{/\mathcal{S}} = \begin{bmatrix} 0 & 0 \\ 0 & W_{22} - d^*d \end{bmatrix},$$

2. *si $\mathcal{M} = \overline{W^{1/2}(\mathcal{S})}$ y $P_{\mathcal{M}}$ es la proyección ortogonal en $\mathcal{M}$ entonces*

$$W_{/\mathcal{S}} = W^{1/2}(I - P_{\mathcal{M}})W^{1/2},$$

3. *$R(W) \cap \mathcal{S}^\perp \subseteq R(W_{/\mathcal{S}}) \subseteq R(W_{/\mathcal{S}}^{1/2}) = R(W^{1/2}) \cap \mathcal{S}^\perp$; en general, las inclusiones son estrictas.*

Para las pruebas de estas afirmaciones, ver [4, Teorema 3], [4, Corolario 1] y [71, Teorema 1.4].

A continuación se probarán las siguientes propiedades del núcleo y el rango de $W_{/\mathcal{S}}$ y $W_{\mathcal{S}}$.

Corolario 5.1.7. *Sean $W \in L(\mathcal{H})^+$ y $\mathcal{S} \subseteq \mathcal{H}$ un subespacio cerrado. Entonces*

1. *$\overline{N(W) + \mathcal{S}} \subseteq N(W_{/\mathcal{S}}) = W^{-1/2}\overline{(W^{1/2}(\mathcal{S}))}$.*
2. *$N(W_{/\mathcal{S}}) = N(W) + \mathcal{S}$ si y sólo si $W^{1/2}(\mathcal{S})$ es cerrado en $R(W^{1/2})$, es decir, $R(W^{1/2}) \cap \overline{(W^{1/2}(\mathcal{S}))} = W^{1/2}(\mathcal{S})$.*

Demostración.

1. Por el Teorema 5.1.6, si $\mathcal{M} = \overline{(W^{1/2}(\mathcal{S}))}$, entonces $W_{/\mathcal{S}} = W^{1/2}(I - P_{\mathcal{M}})W^{1/2}$. Por lo tanto $N(W)$ y $\mathcal{S}$ están incluidos en $N(W_{/\mathcal{S}})$. Por otra parte,
$$N(W_{/\mathcal{S}}) = N(W^{1/2}(I - P_{\mathcal{M}})W^{1/2}) = N(I - P_{\mathcal{M}})W^{1/2}) = W^{-1/2}(\mathcal{M}).$$
2. Si $\mathcal{M} = \overline{(W^{1/2}(\mathcal{S}))}$, es claro que $W^{1/2}(\mathcal{S})$ es cerrado en $R(W^{1/2})$, si y sólo si $W^{1/2}(\mathcal{S}) = \mathcal{M} \cap R(W^{1/2})$, si y sólo si
$$W^{-1/2}(\mathcal{M}) = W^{-1/2}(W^{1/2}(\mathcal{S})) = N(W) + \mathcal{S}.$$

□

Ahora probaremos varias propiedades útiles del operador de shorted.

Proposición 5.1.8 ([4]). *Sean $W \in L(\mathcal{H})^+$ y $\mathcal{S}$, $\mathcal{T}$ subespacios cerrados de $\mathcal{H}$. Entonces*
$$(W_{/\mathcal{T}^\perp})_{/\mathcal{S}^\perp} = W_{/(\mathcal{S}\cap\mathcal{T})^\perp}.$$

Demostración. Sean
$$\mathcal{M}_1 = \mathcal{M}(W, (\mathcal{S} \cap \mathcal{T})^\perp) = \{X \in L(\mathcal{H}) : 0 \leq X \leq W \text{ y } R(X) \subseteq \mathcal{S} \cap \mathcal{T}\},$$
$$\mathcal{M}_2 = \mathcal{M}(W_{/\mathcal{T}^\perp}, \mathcal{S}^\perp) = \{X \in L(\mathcal{H}) : 0 \leq X \leq W_{/\mathcal{T}^\perp} \text{ y } R(X) \subseteq \mathcal{S}\}.$$
Veamos que $\mathcal{M}_1 = \mathcal{M}_2$ y por lo tanto sus máximos coinciden. Si $X \in \mathcal{M}_1$ entonces $R(X) \subseteq \mathcal{S} \cap \mathcal{T} \subseteq \mathcal{T}$ y $0 \leq X \leq W$; entonces $X \leq W_{\mathcal{T}^\perp}$. Más aún, $R(X) \subseteq \mathcal{S}$ y entonces $X \in \mathcal{M}_2$. Recíprocamente, si $X \in \mathcal{M}_2$, entonces $R(X) \subseteq \mathcal{S}$ y $0 \leq X \leq W_{/\mathcal{T}^\perp}$, entonces por el Teorema de Douglas 2.1.1, vale que $R(X^{1/2}) \subseteq R(W_{/\mathcal{T}^\perp}^{1/2})$, entonces
$$R(X) \subseteq R(X^{1/2}) \subseteq R(W_{/\mathcal{T}^\perp}^{1/2}) \subseteq N(W_{/\mathcal{T}^\perp})^\perp \subseteq \mathcal{T},$$

ya que $\mathcal{T}^\perp \subseteq N(W_{/\mathcal{T}^\perp})$. Por lo tanto $R(X) \subseteq \mathcal{S} \cap \mathcal{T}$ y además $0 \leq X \leq W_{/\mathcal{T}^\perp} \leq W$, entonces $X \in \mathcal{M}_1$. □

Corolario 5.1.9. *Sean $W \in L(\mathcal{H})^+$ y $\mathcal{S}, \mathcal{T}$ subespacios cerrados de $\mathcal{H}$. Entonces*

$$(W_{/\mathcal{T}})_{/\mathcal{S}} = W_{/\overline{(\mathcal{S}+\mathcal{T})}}.$$

Demostración. Se sigue de la Proposición 5.1.8 y del hecho que $(\mathcal{S}^\perp \cap \mathcal{T}^\perp)^\perp = \overline{\mathcal{S}+\mathcal{T}}$. □

Proposición 5.1.10. *Sean $W, Y \in L(\mathcal{H})^+$ y $\mathcal{S}$ un subespacio cerrado de $\mathcal{H}$. Entonces*

$$W_{/\mathcal{S}} + Y_{/\mathcal{S}} \leq (W+Y)_{/\mathcal{S}}.$$

Demostración. Tenemos que $0 \leq W_{/\mathcal{S}} + Y_{/\mathcal{S}} \leq W + Y$ y además $R(W_{/\mathcal{S}} + Y_{/\mathcal{S}}) \subseteq \mathcal{S}^\perp$, por lo tanto $W_{/\mathcal{S}} + Y_{/\mathcal{S}} \leq (W+Y)_{/\mathcal{S}}$. □

Observación 5.1.11. Usando el Teorema 5.1.6 y las propiedades probadas para el operador de shorted, es posible deducir fácilmente, las siguientes propiedades de $W_{\mathcal{S}}$:

1. $(W_{\mathcal{S}})_{/\mathcal{S}} = 0$. De hecho, si escribimos $W = W_{/\mathcal{S}} + W_{\mathcal{S}}$. Por la Proposición 5.1.10, tenemos que

$$W_{/\mathcal{S}} = (W_{/\mathcal{S}} + W_{\mathcal{S}})_{/\mathcal{S}} \geq (W_{/\mathcal{S}})_{/\mathcal{S}} + (W_{\mathcal{S}})_{/\mathcal{S}}. \tag{5.4}$$

   Por el Corolario 5.1.9, tenemos que $(W_{/\mathcal{S}})_{/\mathcal{S}} = W_{/\overline{\mathcal{S}+\mathcal{S}}} = W_{/\mathcal{S}}$, y volviendo a (5.4), tenemos que $W_{/\mathcal{S}} \geq W_{/\mathcal{S}} + (W_{\mathcal{S}})_{/\mathcal{S}}$, entonces $0 \geq (W_{\mathcal{S}})_{/\mathcal{S}} \geq 0$, y por lo tanto $(W_{\mathcal{S}})_{/\mathcal{S}} = 0$,

2. Por el Teorema 5.1.4, como $W \in L(\mathcal{H})^+$, tenemos que $R(W_{12}) \subseteq R(W_{11}^{1/2})$. Sea $d \in L(\mathcal{S}^\perp, \mathcal{S})$ la solución reducida de la ecuación $W_{11}^{1/2}x = W_{12}$, entonces

$$W_{\mathcal{S}} = \begin{bmatrix} W_{11} & W_{12} \\ W_{12}^* & d^*d \end{bmatrix} = \begin{bmatrix} W_{11}^{1/2} & 0 \\ d^* & 0 \end{bmatrix} \begin{bmatrix} W_{11}^{1/2} & d \\ 0 & 0 \end{bmatrix},$$

3. $N(W_{\mathcal{S}}) = W^{-1}(\mathcal{S}^\perp)$. Si $\mathcal{M} = \overline{W^{1/2}(\mathcal{S})}$, tenemos que $\mathcal{M}^\perp = W^{-1/2}(\mathcal{S}^\perp)$, entonces

$$\begin{aligned} N(W_{\mathcal{S}}) &= N(P_{\mathcal{M}} W^{1/2}) = W^{-1/2}(\mathcal{M}^\perp) = \\ &= W^{-1/2}(W^{-1/2}(\mathcal{S}^\perp)) = W^{-1}(\mathcal{S}^\perp), \end{aligned}$$

4. $W(\mathcal{S}) \subseteq R(W_{\mathcal{S}}) \subseteq \overline{W(\mathcal{S})}$ y las inclusiones pueden ser estrictas. De hecho, como $\mathcal{S} \subseteq N(W_{/\mathcal{S}})$,

$$W(\mathcal{S}) = W_{\mathcal{S}}(\mathcal{S}) \subseteq R(W_{\mathcal{S}}) \subseteq N(W_{\mathcal{S}})^\perp = (W^{-1}(\mathcal{S}^\perp))^\perp = \overline{W(\mathcal{S})}.$$

5.2 OPERADOR DE SHORTED Y COMPATIBILIDAD

La siguiente proposición relaciona el operador shorted $W_{/\mathcal{S}}$ con los elementos de $\mathcal{P}(W,\mathcal{S})$, cuando el par $(W,\mathcal{S})$ es compatible.

Proposición 5.2.1 ([23]). *Sea $W \in L(\mathcal{H})^+$ tal que el par $(W,\mathcal{S})$ es compatible. Sea $E \in \mathcal{P}(W,\mathcal{S})$ y $Q = I - E$. Entonces*

1. $W_{/\mathcal{S}} = WQ = Q^*WQ$,
2. $W_{/\mathcal{S}} = \text{mín}\,\{E^*WE : E \in \mathcal{Q},\ N(E) = \mathcal{S}\}$. *De hecho, esta propiedad es equivalente a la compatibilidad del par $(W,\mathcal{S})$,*
3. $R(W_{/\mathcal{S}}) = R(W) \cap \mathcal{S}^\perp$,
4. $N(W_{/\mathcal{S}}) = N(W) + \mathcal{S}$.

Demostración.

1. Notar que $0 \leq WQ = Q^*WQ \leq W$, por la Proposición 4.1.1. También $R(WQ) = R(Q^*W) \subseteq R(Q^*) = R(I - E^*) = N(E^*) = R(E)^\perp = \mathcal{S}^\perp$; entonces $Q^*WQ \leq W_{/\mathcal{S}}$. Dado $0 \leq X \leq W$ con $R(X) \subseteq \mathcal{S}^\perp$, ya que $N(Q) = \mathcal{S}$, tenemos que
$$X = Q^*XQ \leq Q^*WQ = WQ,$$
donde la primera igualdad se puede chequear fácilmente porque, usando la descomposición presentanda en (5.1), X tiene la forma $\begin{bmatrix} 0 & 0 \\ 0 & x \end{bmatrix}$, con $x \in L(\mathcal{S}^\perp,\mathcal{S}^\perp)$ y Q^* tiene la forma $\begin{bmatrix} 0 & 0 \\ z & I \end{bmatrix}$, con $z \in L(\mathcal{S},\mathcal{S}^\perp)$.

2. Por el item 1., $Q^*WQ = W_{/\mathcal{S}}$ y $N(Q) = \mathcal{S}$. Entonces el mínimo se alcanza en Q por el Teorema 5.1.5.

   Recíprocamente, si el mínimo se alcanza para alguna proyección E con $N(E) = \mathcal{S}$, entonces $E^*WE = W_{/\mathcal{S}} \leq W$, implica que E es W-autoadjunto por la Proposición 4.1.1. Entonces $I - E \in \mathcal{P}(W,\mathcal{S})$.

3. Claramente la ecuación $W_{/\mathcal{S}} = WQ = Q^*W$ implica que $R(W_{/\mathcal{S}}) \subseteq R(W) \cap \mathcal{S}^\perp$. La otra inclusión siempre vale por el Teorema 5.1.6.

4. Como el par $(W,\mathcal{S})$ es compatible, tenemos que $\mathcal{H} = \mathcal{S} + \mathcal{S}^{\perp_W}$ (ver Lema 4.1.2), por lo tanto aplicando $W^{1/2}$ a ambos lados de la igualdad obtenemos que
$$W^{1/2}(\mathcal{H}) = W^{1/2}(\mathcal{S}) + W^{1/2}(W^{-1}(\mathcal{S}^\perp)),$$

ó equivalentemente

$$R(W^{1/2}) = W^{1/2}(\mathcal{S}) + W^{-1/2}(\mathcal{S}^{\perp}) \cap R(W^{1/2}) =$$

$$= W^{1/2}(\mathcal{S}) \oplus W^{1/2}(\mathcal{S})^{\perp} \cap R(W^{1/2}).$$

En este caso, si $\mathcal{M} = \overline{W^{1/2}(\mathcal{S})}$ entonces

$$R(W^{1/2}) = \mathcal{M} \cap R(W^{1/2}) \oplus \mathcal{M}^{\perp} \cap R(W^{1/2}) =$$

$$W^{1/2}(\mathcal{S}) \oplus \mathcal{M}^{\perp} \cap R(W^{1/2}),$$

por lo tanto $W^{1/2}(\mathcal{S}) = \mathcal{M} \cap R(W^{1/2})$, entonces $W^{1/2}(\mathcal{S})$ es cerrado en $R(W^{1/2})$, y por el Corolario 5.1.7 $N(W_{/\mathcal{S}}) = N(W) + \mathcal{S}$.

□

La condición $R(W_{/\mathcal{S}}) \subseteq R(W)$, la cual es necesaria para la compatibilidad del par $(W, \mathcal{S})$, implica que algún subespacio más grande que $\mathcal{S}$ (de hecho $N(W_{/\mathcal{S}})$) es W-compatible. La siguiente proposición, probada en [23] muestra exactamente eso.

Proposición 5.2.2 ([23]). *Sea $W \in L(\mathcal{H})^+$ y $\mathcal{S} \subseteq \mathcal{H}$ un subespacio cerrado tal que $R(W_{/\mathcal{S}}) \subseteq R(W)$. Denotemos $\mathcal{T} = N(W_{/\mathcal{S}})$. Entonces el par $(W, \mathcal{T})$ es compatible.*

Demostración. La condición $R(W_{/\mathcal{S}}) \subseteq R(W)$ implica, por el Teorema de Douglas 2.1.1 que el conjunto

$$\Delta = \{Q \in L(\mathcal{H}) : WQ = W_{/\mathcal{S}} \text{ y } N(Q) = \mathcal{T}\}$$

no es vacío. Sea $Q \in \Delta$, entonces Q verifica que $N(Q) = \mathcal{T}$ y $Q^*W = WQ$, pues $W_{/\mathcal{S}}$ es autoadjunto, sólo resta probar que $Q^2 = Q$.

Primero probemos que si $\mathcal{Z} = W^{-1/2}(\mathcal{S}^{\perp}) = W^{1/2}(\mathcal{S})^{\perp}$, entonces Q es una solución de $W^{1/2}X = P_{\mathcal{Z}}W^{1/2}$. Recordar que, por el Teorema 5.1.6, $W_{/\mathcal{S}} = W^{1/2}P_{\mathcal{Z}}W^{1/2}$, entonces

$$W^{1/2}(W^{1/2}Q - P_{\mathcal{Z}}W^{1/2}) = 0.$$

Si $\xi \in \mathcal{H}$, $P_{\mathcal{Z}}W^{1/2}\xi = W^{1/2}Q\xi + \alpha$, con $\alpha \in N(W^{1/2}) = R(W^{1/2})^{\perp} \subseteq \mathcal{Z}$. Entonces

$$\|\alpha\|^2 = \langle P_{\mathcal{Z}}W^{1/2}\xi, \alpha\rangle - \langle W^{1/2}Q\xi, \alpha\rangle = \langle W^{1/2}\xi, P_{\mathcal{Z}}\alpha\rangle = \langle W^{1/2}\xi, \alpha\rangle = 0,$$

pues $\alpha \in N(W^{1/2})$ y $P_{\mathcal{Z}}\alpha = \alpha$.

Por lo tanto $W^{1/2}Q = P_{\mathcal{Z}}W^{1/2}$.

Notar también que $W^{1/2}Q^2 = (P_{\mathcal{Z}})^2W^{1/2} = P_{\mathcal{Z}}W^{1/2}$, entonces

$$W^{1/2}(Q^2 - Q) = 0.$$

Sea $\rho \in R(Q)$. Entonces $Q\rho - \rho \in N(W) \cap R(Q)$. Si $Q\rho - \rho = Q\omega$, para algún $\omega \in \mathcal{H}$, entonces $0 = WQ\omega = W_{/\mathcal{S}}\omega$. Entonces $\omega \in N(W_{/\mathcal{S}}) = \mathcal{T} = N(Q)$. Entonces $Q\rho = \rho$ para todo $\rho \in R(Q)$. Esto implica que $Q^2 = Q$ y por lo tanto el par $(W, \mathcal{T})$ es compatible. □

Usando la proposición anterior, es posible demostrar la siguiente propiedad, la cual nos permite establecer una equivalencia entre la compatibilidad del par $(W, \mathcal{S})$ y el rango y el núcleo del operador shorted $W_{/\mathcal{S}}$.

Teorema 5.2.3 ([23])**.** *Sea $W \in L(\mathcal{H})^+$ y $\mathcal{S} \subseteq \mathcal{H}$ un subespacio cerrado. Entonces el par $(W, \mathcal{S})$ es compatble si y sólo si $R(W_{/\mathcal{S}}) = R(W) \cap \mathcal{S}^\perp$ y $N(W_{/\mathcal{S}}) = N(W) + \mathcal{S}$.*

Demostración. Si $R(W_{/\mathcal{S}}) = R(W) \cap \mathcal{S}^\perp$ y $N(W_{/\mathcal{S}}) = N(W) + \mathcal{S} = \mathcal{T}$ entonces, por la Proposición 5.2.2, el par $(W, \mathcal{T})$ es compatible, o equivalentemente $\mathcal{T} + W^{-1}(\mathcal{T}) = \mathcal{H}$. Pero

$$N(W) \subseteq W^{-1}(\mathcal{S}^\perp) = W(\mathcal{S})^\perp = W(\mathcal{T})^\perp = W^{-1}(\mathcal{T}^\perp),$$

entonces $\mathcal{H} = \mathcal{T} + W^{-1}(\mathcal{T}^\perp) = N(W) + \mathcal{S} + W^{-1}(\mathcal{T}^\perp) = \mathcal{S} + W^{-1}(\mathcal{S}^\perp)$ y el par $(W, \mathcal{S})$ es compatible. La recíproca fue probada en la Proposición 5.2.1. □

6 PROBLEMAS DE CUADRADOS MÍNIMOS CON OPERADORES

6.1 SOLUCIÓN DE CUADRADOS MÍNIMOS CON PESO DE LA ECUACIÓN AX-B$=0$

Dado $W \in L(\mathcal{H})^+$, tal que $W^{1/2} \in S_p$ para algún p con $1 \leq p < \infty$, consideremos la seminorma de operadores asociada a W,

$$\|X\|_{p,W} = \|W^{1/2}X\|_p,$$

para $X \in L(\mathcal{H})$. Estudiamos el siguiente problema de aproximación: dados $A \in CR(\mathcal{H})$ y $B \in L(\mathcal{H})$, analizar la existencia de

$$\underset{X\in L(\mathcal{H})}{min} \|AX - B\|_{p,W}.$$

Por simplicidad, en esta sección, estudiamos primero el caso $B = I$, es decir, estudiamos el problema

$$\underset{X\in L(\mathcal{H})}{min} \|AX - I\|_{p,W}. \tag{6.1}$$

Para estudiar el Problema (6.1) primero introducimos el siguiente problema asociado: dados $W \in L(\mathcal{H})^+$, $A \in CR(\mathcal{H})$ y $F : L(\mathcal{H}) \to L(\mathcal{H})^+$,

$$F(X) = (AX - I)^*W(AX - I),$$

analizamos,

$$\underset{X\in L(\mathcal{H})}{inf} \ F(X), \tag{6.2}$$

con el orden inducido en $L(\mathcal{H})$ por el cono de operadores positivos. Los resultados presentados en esta sección fueron publicados en [18].

La siguiente proposición muestra que el ínfimo de la ecuación (6.2) siempre existe y coincide con el operador de shorted de W en $R(A)$.

6.1.1 *Solución de cuadrados mínimos con peso de la ecuación $AX - I = 0$*

Proposición 6.1.1. *Sea $A \in CR(\mathcal{H})$ y $W \in L(\mathcal{H})^+$, entonces el ínfimo del Problema (6.2) existe y además*

$$\underset{X\in L(\mathcal{H})}{inf} \ F(X) = W_{/R(A)}.$$

Demostración. Sea $X \in L(\mathcal{H})$, escribiendo $W = W_{/R(A)} + W_{R(A)}$, se sigue que

$$(I - AX)^* W (I - AX) =$$

$$= (I - AX)^* W_{/R(A)} (I - AX) + (I - AX)^* W_{R(A)} (I - AX)$$
$$= W_{/R(A)} + (I - AX)^* W_{R(A)} (I - AX) \geq W_{/R(A)},$$

porque $R(A) \subseteq N(W_{/R(A)})$ (ver Corolario 5.1.7) y entonces $W_{/R(A)}(I - AX) = W_{/R(A)} = (I - AX)^* W_{/R(A)}$. Por lo tanto $W_{/R(A)}$ es una cota inferior de $F(X)$.

Si $C \geq 0$ es otra cota inferior de $F(X)$, entonces

$$C \leq F(X), \text{ para cada } X \in L(\mathcal{H}).$$

En particular,

$$C \leq E^* W E,$$

donde E es cualquier proyección tal que $N(E) = R(A)$. De hecho $R(I - E) = N(E) = R(A)$, entonces por el Teorema de Douglas 2.1.1, existe $X_0 \in L(\mathcal{H})$, tal que $(I - E) = AX_0$, es decir, $(-E) = AX_0 - I$. Por lo tanto, por el Teorema 5.1.5

$$C \leq inf\{E^* W E : \ E^2 = E, \ N(E) = R(A)\} = W_{/R(A)}.$$

Entonces,

$$W_{/R(A)} = \underset{X \in L(\mathcal{H})}{inf} F(X).$$

□

Teorema 6.1.2. *Sea $A \in CR(\mathcal{H})$ y $W \in L(\mathcal{H})^+$. Entonces el Problema (6.2) tiene mínimo, es decir, existe $X_0 \in L(\mathcal{H})$ tal que*

$$F(X_0) = \underset{X \in L(\mathcal{H})}{min} F(X) = W_{/R(A)}$$

si y sólo si el par $(W, R(A))$ es compatible.

Demostración. Supongamos que el Problema (6.2) tiene un mínimo, en este caso, por la Proposición 6.1.1 se cumple que

$$\underset{X \in L(\mathcal{H})}{min} F(X) = W_{/R(A)}.$$

Si $X_0 \in L(\mathcal{H})$ es tal que $F(X_0) = \underset{X \in L(\mathcal{H})}{min} F(X) = W_{/R(A)}$, entonces dado que $W = W_{/R(A)} + W_{R(A)}$, se sigue que

$$W_{/R(A)} = (AX_0 - I)^* W (AX_0 - I) = W_{/R(A)} + (AX_0 - I)^* W_{R(A)} (AX_0 - I).$$

Por lo tanto

$$(AX_0 - I)^* W_{R(A)} (AX_0 - I) = 0,$$

entonces,

$$W_{R(A)}^{1/2}(AX_0 - I) = 0,$$

y entonces por la Observación 5.1.11

$$R(AX_0 - I) \subseteq N(W_{R(A)}) = W^{-1}(R(A)^{\perp}).$$

Por lo tanto

$$W(R(AX_0 - I)) \subseteq R(A)^{\perp} \cap R(W).$$

Entonces

$$R(W_{/R(A)}) = R((AX_0 - I)^* W(AX_0 - I)) \subseteq (AX_0 - I)^*(R(A)^{\perp} \cap R(W)) =$$

$$R(A)^{\perp} \cap R(W),$$

porque $A^*(R(A)^{\perp}) = \{0\}$. Entonces, $R(W_{/R(A)}) = R(A)^{\perp} \cap R(W)$, pues $R(A)^{\perp} \cap R(W)$ siempre está contenido in $R(W_{/R(A)})$ (ver Teorema 5.1.6).

También $x \in N(W_{/R(A)})$ si y sólo si $W^{1/2}(AX_0 - I)x = 0$, ó equivalentemente $(AX_0 - I)x \in N(W)$. En este caso $x \in N(W) + R(A)$, y entonces

$$N(W_{/R(A)}) = N(W) + R(A),$$

pues $N(W) + R(A)$ está siempre contenido en $N(W_{/R(A)})$ (ver Corolario 5.1.7). Por lo tanto $R(W_{/R(A)}) = R(A)^{\perp} \cap R(W)$ y $N(W_{/R(A)}) = N(W) + R(A)$ y por el Teorema 5.2.3, el par $(W, R(A))$ es compatible.

Recíprocamente, si el par $(W, R(A))$ es compatible, entonces por la Proposición 5.2.1,

$$W_{/R(A)} = min\{E^*WE : \ E^2 = E, \ N(E) = R(A)\}.$$

Sea E_0 tal que $E_0^2 = E_0$, $N(E_0) = R(A)$ y $W_{/R(A)} = E_0^* W E_0$. Considerar $X_0 = A^{\dagger}(I - E_0)$, entonces $-E_0 = AX_0 - I$ y $F(X_0) = W_{/R(A)}$. □

Corolario 6.1.3. *Sea $A \in CR(\mathcal{H})$ y $W \in L(\mathcal{H})^+$. Entonces el Problema* (6.2) *tiene mínimo, es decir, existe $X_0 \in L(\mathcal{H})$ tal que*

$$F(X_0) = \min_{X \in L(\mathcal{H})} F(X) = W_{/R(A)}$$

si y sólo si el ángulo de Dixmier $c_0(R(A)^{\perp}, \overline{W(R(A))}) < 1$.

Demostración. Se sigue del Teorema 6.1.2 y del Teorema 4.1.5. □

Si el par $(W, R(A))$, es compatible entonces por el Teorema 6.1.2, el Problema (6.2) alcanza un mínimo, es decir, existe $X_0 \in L(\mathcal{H})$ tal que $F(X_0) = W_{/R(A)}$. Considerar el conjunto

$$M = \{X \in L(\mathcal{H}) : F(X) = W_{/R(A)}\}.$$

La próxima proposición da una caracterización de los elementos de M.

Proposición 6.1.4. *Si el par $(W, R(A))$ es compatible, entonces las siguientes condiciones son equivalentes:*

1. *$X_0 \in M$, es decir $F(X_0) = \min\limits_{X \in L(\mathcal{H})} F(X)$,*
2. *X_0 es una W-inversa de A,*
3. *X_0 es una solución de la ecuación normal*

$$A^*W(AX - I) = 0.$$

Demostración. 1. $\Leftrightarrow$ 2. : Sea $X_0 \in M$ entonces

$$\langle F(X_0)x, x \rangle \leq \langle F(X)x, x \rangle, \text{ para cada } x \in \mathcal{H} \text{ y cada } X \in L(\mathcal{H}),$$

ó equivalentemente

$$\|AX_0x - x\|_W^2 \leq \|AXx - x\|_W^2, \text{ para cada } x \in \mathcal{H} \text{ y cada } X \in L(\mathcal{H}).$$

Para cada $x \in \mathcal{H}$, dado $z \in \mathcal{H}$, sea $X \in L(\mathcal{H})$ tal que $z = Xx$. Entonces

$$\|AX_0x - x\|_W^2 \leq \|Az - x\|_W^2, \text{ para cada } x \in \mathcal{H} \text{ y cada } z \in \mathcal{H}.$$

Por lo tanto X_0 es una W-inversa de A.

Recíprocamente, dado que el par $(W, R(A))$ es compatible, por el Teorema 4.2.3, el operador A admite una W-inversa. Si X_0 es una W-inversa de A, entonces dado $X \in L(\mathcal{H})$ obtenemos

$$\|AX_0x - x\|_W^2 \leq \|AXx - x\|_W^2, \text{ para cada } x \in \mathcal{H},$$

entonces $F(X_0) \leq F(X)$ para cada $X \in L(\mathcal{H})$, es decir $X_0 \in M$.

La equivalencia 2. $\Leftrightarrow$ 3. fue establecida en el Teorema 4.2.3, para $B = I$. □

El siguiente resultado prueba la equivalencia entre la existencia del mínimo del Problema (6.1) y la compatibilidad del par $(W, R(A))$.

Teorema 6.1.5. *Sea $A \in CR(\mathcal{H})$ y $W \in L(\mathcal{H})^+$, tal que $W^{1/2} \in S_p$, para algún p con $1 \leq p < \infty$. El Problema (6.1) tiene mínimo si y sólo si el par $(W, R(A))$ es compatible.*

En este caso,

$$\min_{X \in L(\mathcal{H})} \|AX - I\|_{p,W} = \|W^{1/2}_{/R(A)}\|_p.$$

Mas aún, $X_0 \in L(\mathcal{H})$ satisface

$$\|AX_0 - I\|_{p,W} = \|W^{1/2}_{/R(A)}\|_p,$$

si y sólo si X_0 es una W-inversa de A.

Demostración. Si el par $(W, R(A))$ es compatible, entonces por el Teorema 6.1.2, existe $X_0 \in L(\mathcal{H})$ tal que $F(X_0) = min_{X \in L(\mathcal{H})} F(X) = W_{/R(A)}$, es decir

$$W_{/R(A)} = F(X_0) \leq F(X), \text{ para cada } X \in L(\mathcal{H}).$$

Dado que $W^{1/2} \in S_p$, por la Proposición 3.1.15,

$$\|W^{1/2}_{/R(A)}\|_p = \|W^{1/2}(AX_0 - I)\|_p = \|AX_0 - I\|_{p,W} \leq \|AX - I\|_{p,W},$$

para cada $X \in L(\mathcal{H})$, entonces

$$\underset{X \in L(\mathcal{H})}{min} \|AX - I\|_{p,W} = \|AX_0 - I\|_{p,W} = \|W^{1/2}_{/R(A)}\|_p.$$

Para probar la recíproca, para $1 \leq p < \infty$, consideremos $F_p : S_p \to \mathbb{R}^+$,

$$F_p(X) = \|W^{1/2}(AX - I)\|_p^p.$$

Por el Teorema 3.2.7 y el Teorema 3.2.10, F_p tiene una derivada $\phi - direccional$ para todo $\phi \in [0, 2\pi)$. Entonces es fácil chequear que, para cada $X, Y \in L(\mathcal{H})$ y $\phi \in [0, 2\pi)$,

$$D_\phi F_p(X, Y) = D_\phi G_p(W^{1/2}(AX - I), W^{1/2}AY),$$

donde $G_p(X) = \|X\|_p^p$.

Supongamos que el Problema (6.1) admite mínimo, es decir existe $X_0 \in L(\mathcal{H})$, un mínimo global de $\|AX - I\|_{p,W}$. Entonces X_0 es un mínimo global de F_p y, por Lema 3.2.4, tenemos

$$\underset{0 \leq \phi < 2\pi}{inf} (D_\phi F_p(X_0, Y)) \geq 0, \text{ para cada } Y \in L(\mathcal{H}).$$

Sea $W^{1/2}(AX_0 - I) = U|W^{1/2}(AX_0 - I)|$ la descomposición polar del operador $W^{1/2}(AX_0 - I)$, con U una isometría parcial con $N(U) = N(W^{1/2}(AX_0 - I))$, $P = P_{N(W^{1/2}(AX_0 - I))}$ y $Q = P_{N((W^{1/2}(AX_0 - I))^*)}$.

Si $p = 1$, por el Corolario 3.2.14 se sigue, para cada $\phi \in [0, 2\pi)$

$$0 \leq D_\phi F_1(X_0, Y) = Re\,[e^{i\phi} tr(U^* W^{1/2} AY)] + \|QW^{1/2}AYP\|_1,$$

para cada $Y \in L(\mathcal{H})$. Considerando una ϕ adecuada para cada $Y \in L(\mathcal{H})$, obtenemos

$$|tr(U^* W^{1/2} AY)| \leq \|QW^{1/2}AYP\|_1, \text{ para cada } Y \in L(\mathcal{H}).$$

Observar que $R(Q) = N(U^*)$ y $R(P) = N(U)$, por lo tanto $U^*Q = PU^* = 0$.

Sea $Y \in L(\mathcal{H})$ entonces $|tr(U^* W^{1/2} AY)| = |tr((I - P)U^* W^{1/2} AY)| = |tr(U^* W^{1/2} AY(I - P))| \leq \|QW^{1/2}AY(I - P)P\|_1 = 0$. Entonces

$$tr(U^* W^{1/2} AY) = 0, \text{ para cada } Y \in L(\mathcal{H}).$$

Por lo tanto
$$U^*W^{1/2}A = 0.$$
Entonces
$$R(W^{1/2}A) \subseteq N(U^*) = N((AX_0 - I)^*W^{1/2}).$$
Por lo tanto
$$(AX_0 - I)^*W^{1/2}(W^{1/2}A) = 0,$$
ó equivalentemente
$$A^*WAX_0 = A^*W,$$
y por el Teorema 4.2.3 y el Teorema 4.1.3, el par $(W, R(A))$ es compatible.

Si $1 < p < \infty$, por el Corolario 3.2.8 se sigue, para cada $\phi \in [0, 2\pi)$
$$0 \leq D_\phi F_p(X_0, Y) = p\ Re\ [e^{i\phi} tr(|W^{1/2}(AX_0 - I)|^{p-1}U^*W^{1/2}AY)],$$
para cada $Y \in L(\mathcal{H})$. Tomando ϕ adecuado para cada Y, se sigue que
$$|W^{1/2}(AX_0 - I)|^{p-1}U^*W^{1/2}A = 0.$$
Por la Proposición A.2.3, vale que
$N(|W^{1/2}(AX_0 - I)|^{p-1}) = N(|W^{1/2}(AX_0 - I)|)$, entonces se sigue que
$$|W^{1/2}(AX_0 - I)|U^*W^{1/2}A = 0,$$
y por lo tanto
$$A^*WAX_0 = A^*W.$$
Por el Teorema 4.2.3 y el Teorema 4.1.3, el par $(W, R(A))$ es compatible.

Finalmente, si $X_0 \in L(\mathcal{H})$ minimiza el Problema (6.1), hemos probado que X_0 es una solución de la ecuación normal $A^*W(AX - I) = 0$ y por la Proposición 6.1.4, es una W- inversa de A. Recíprocamente, si X_0 es una W-inversa de A, entonces por la Proposición 6.1.4, X_0 minimiza (6.2), y por la Proposición 3.1.15, minimiza (6.1). □

Observación 6.1.6. Sea $|||\cdot|||$ una norma unitariamente invariante en un ideal no nulo $\mathcal{J}$ de $L(\mathcal{H})$. Dado $W \in L(\mathcal{H})^+$ tal que $W^{1/2} \in \mathcal{J}$, considerar la seminorma asociada a W dada por
$$\|X\|_W = |||W^{1/2}X|||,$$
para $X \in L(\mathcal{H})$. Sea $A \in CR(\mathcal{H})$, si el par $(W, R(A))$ es compatible, entonces existe el mínimo del conjunto $\{\|AX - I\|_W : X \in L(\mathcal{H})\}$ y
$$\min_{X \in L(\mathcal{H})} \|AX - I\|_W = |||W^{1/2}_{/R(A)}|||.$$
En particular si $\mathcal{J} = L(\mathcal{H})$ y consideramos la norma de operadores $\|\cdot\|$, la observación también vale.

De hecho, si el par $(W, R(A))$ es compatible, por el Teorema 6.1.2, existe $X_0 \in L(\mathcal{H})$ tal que $F(X_0) = min_{X \in L(\mathcal{H})} F(X) = W_{/R(A)}$, es decir

$$W_{/R(A)} = F(X_0) \leq F(X), \text{ para cada } X \in L(\mathcal{H}).$$

Dado que $W^{1/2} \in \mathcal{J}$, por Proposición 3.1.15,

$$|||W^{1/2}_{/R(A)}||| = |||W^{1/2}(AX_0 - I)||| \leq |||W^{1/2}(AX - I)||| = \|AX - I\|_W,$$

para cada $X \in L(\mathcal{H})$, entonces

$$\min_{X \in L(\mathcal{H})} \|AX - I\|_W = \|AX_0 - I\|_W = |||W^{1/2}_{/R(A)}|||.$$

6.1.2 *Solución de cuadrados mínimos con peso de la ecuación $AX - B = 0$*

En esta sección, estudiamos el problema general presentado al principio del capítulo: dado $A \in CR(\mathcal{H})$, $B \in L(\mathcal{H})$ y $W \in L(\mathcal{H})^+$ tal que $W^{1/2} \in S_p$ para algún p con $1 \leq p < \infty$, analizar la existencia de

$$\min_{X \in L(\mathcal{H})} \|AX - B\|_{p,W}. \tag{6.3}$$

Para estudiar el Problema (6.3) introducimos el siguiente problema asociado: dado $A \in CR(\mathcal{H})$, $B \in L(\mathcal{H})$, $W \in L(\mathcal{H})^+$ y $G : L(\mathcal{H}) \to L(\mathcal{H})$,

$$G(X) = (AX - B)^* W (AX - B),$$

analizar la existencia de

$$\inf_{X \in L(\mathcal{H})} G(X), \tag{6.4}$$

Lema 6.1.7. *Sea $A \in CR(\mathcal{H})$, $B \in L(\mathcal{H})$ y $W \in L(\mathcal{H})^+$, entonces el conjunto $\{B^*E^*WEB : \ E^2 = E, \ N(E) = R(A)\}$ tiene ínfimo y*

$$B^* W_{/R(A)} B = inf\{B^*E^*WEB : \ E^2 = E, \ N(E) = R(A)\}.$$

Demostración. Si W es inversible, entonces el par $(W, R(A))$ es compatible (ver 4.1) y, por la Proposición 5.2.1,

$$W_{/R(A)} = \min \{E^*WE : E^2 = E, \ N(E) = R(A)\}.$$

En este caso existe una única proyección E_0 donde se alcanza el mínimo (ver [24]), es decir $W_{/R(A)} = E_0^* W E_0$. Entonces

$$B^* W_{/R(A)} B = B^* E_0^* W E_0 B \leq B^* E^* W E B,$$

para cada proyección E con $N(E) = R(A)$, y entonces

$$min\{B^*E^*WEB : \; E^2 = E, \; N(E) = R(A)\} = B^*W_{/R(A)}B.$$

Para $W \in L(\mathcal{H})^+$ no inversibles, por el Teorema 5.1.5, siempre vale que

$$B^*W_{/R(A)}B \leq B^*E^*WEB, \text{ para cada proyección } E \text{ tal que } N(E) = R(A).$$

Por lo tanto $B^*W_{/R(A)}B$ es una cota inferior de $\{B^*E^*WEB : \; E^2 = E, \; N(E) = R(A)\}$.

Si $C \geq 0$ es otra cota inferior de $\{B^*E^*WEB : \; E^2 = E, \; N(E) = R(A)\}$, entonces para cualquier $\varepsilon > 0$, y cualquier proyeccíon $E \in L(\mathcal{H})$ con $N(E) = R(A)$, tenemos

$$C \leq B^*E^*WEB \leq B^*E^*(W + \varepsilon I)EB.$$

Dado que $W + \varepsilon I$ es positivo e inversible, se sigue que $C \leq B^*(W + \varepsilon I)_{/R(A)}B$, y dado que ε is arbitrario, por el Corolario 5.1.3, concluimos que $C \leq B^*W_{/R(A)}B$.

□

Proposición 6.1.8. *Sea $A \in CR(\mathcal{H})$, $B \in L(\mathcal{H})$ y $W \in L(\mathcal{H})^+$, entonces el ínfimo del Problema (6.4) existe y además*

$$\underset{X \in L(\mathcal{H})}{inf} \; G(X) = B^*W_{/R(A)}B.$$

Demostración. Siguiendo la misma idea que en la Proposición 6.1.1, sea $X \in L(\mathcal{H})$, entonces

$$(B - AX)^*W(B - AX) =$$
$$(B - AX)^*W_{/R(A)}(B - AX) + (B - AX)^*W_{R(A)}(B - AX) =$$
$$B^*W_{/R(A)}B + (B - AX)^*W_{R(A)}(B - AX) \geq B^*W_{/R(A)}B,$$

pues $R(A) \subseteq N(W_{/R(A)})$. Entonces $B^*W_{/R(A)}B$ es una cota inferior de $G(X)$. Si $C \geq 0$ es otra cota inferior de $G(X)$, entonces

$$C \leq G(X), \text{ para cada } X \in L(\mathcal{H}).$$

En particular,

$$C \leq B^*E^*WEB,$$

donde E es cualquier proyección tal que $N(E) = R(A)$; de hecho $R((I - E)B) \subseteq R(I - E) = N(E) = R(A)$, entonces por el Teorema 2.1.1, existe $X_0 \in L(\mathcal{H})$, tal que $(I - E)B = AX_0$, es decir $(-EB) = AX_0 - B$.

Por lo tanto por el Lema 6.1.7

$$C \leq inf\{B^*E^*WEB : \; E^2 = E, \; N(E) = R(A)\} = B^*W_{/R(A)}B.$$

Entonces,

$$B^*W_{/R(A)}B = \underset{X \in L(\mathcal{H})}{inf} \; G(X).$$

□

Teorema 6.1.9. *Sea $A \in CR(\mathcal{H})$, $W \in L(\mathcal{H})^+$ y $B \in L(\mathcal{H})$. El Problema (6.4) tiene mínimo, es decir, existe $X_0 \in L(\mathcal{H})$ tal que*

$$\min_{X \in L(\mathcal{H})} G(X) = G(X_0) = B^* W_{/R(A)} B$$

si y sólo si $R(B) \subseteq R(A) + R(A)^{\perp_W}$.

Demostración. Supongamos que el Problema (6.4) tiene un mínimo y sea $y \in R(B)$, entonces existe $x \in \mathcal{H}$ tal que $y = Bx$. Si $X_0 \in L(\mathcal{H})$ es tal que $G(X_0) = min_{X \in L(\mathcal{H})} G(X)$ entonces

$$\langle G(X_0)x, x \rangle \leq \langle G(X)x, x \rangle, \text{ para cada } X \in L(\mathcal{H}),$$

ó equivalentemente,

$$\|(AX_0 - B)x\|_W \leq \|(AX - B)x\|_W, \text{ para cada } X \in L(\mathcal{H}).$$

Sea $u_0 = X_0 x$ y sea $z \in \mathcal{H}$ arbitrario, entonces existe $X \in L(\mathcal{H})$ tal que $z = Xx$. Por lo tanto,

$$\|Au_0 - Bx\|_W \leq \|Az - Bx\|_W, \text{ para cada } z \in \mathcal{H}.$$

Por lo tanto u_0 es una $W-$LSS de $Az = Bx$, entonces por el Teorema 4.2.1

$$y = Bx \in R(A) + R(A)^{\perp_W},$$

concluyendo que $R(B) \subseteq R(A) + R(A)^{\perp_W}$.

Recíprocamente, si $R(B) \subseteq R(A) + R(A)^{\perp_W}$, por el Teorema 4.2.3, el operador A admite una W-inversa en $R(B)$. Sea X_0 una W-inversa de A en $R(B)$, entonces

$$\|AX_0 x - Bx\|_W \leq \|Az - Bx\|_W, \text{ para cada } x, z \in \mathcal{H}.$$

En particular, dado $X \in L(\mathcal{H})$, consideremos $z = Xx$. Entonces para cada $x \in \mathcal{H}$,

$$\|AX_0 x - Bx\|_W \leq \|AXx - Bx\|_W.$$

Entonces,

$$\|AX_0 x - Bx\|_W \leq \|AXx - Bx\|_W, \text{para cada } x \in \mathcal{H} \text{ y para cada } X \in L(\mathcal{H}),$$

ó equivalentemente

$$G(X_0) \leq G(X), \text{ para cada } X \in L(\mathcal{H}),$$

por lo tanto el conjunto $\{G(X) : X \in L(\mathcal{H})\}$ admite un elemento mínimo. □

Si $R(B) \subseteq R(A) + R(A)^{\perp_W}$, entonces por el Teorema 6.1.9, el Problema (6.4) alcanza un mínimo, más precisamente probamos que cada W-inversa de A en $R(B)$ minimiza $G(X)$, es decir si $V_0 \in L(\mathcal{H})$ es una W-inversa de A en $R(B)$, entonces $G(V_0) = B^* W_{/R(A)} B$. Consideremos el conjunto

$$M_B = \{X \in L(\mathcal{H}) : G(X) = B^* W_{/R(A)} B\}.$$

La próxima proposición da una caracterización de los elementos de M_B.

Proposición 6.1.10. *Sea* $A \in CR(\mathcal{H})$, $B \in L(\mathcal{H})$ *y* $W \in L(\mathcal{H})^+$. *Si* $R(B) \subseteq R(A) + R(A)^{\perp_W}$ *entonces las siguientes condiciones son equivalentes:*

1. $V \in M_B$, *es decir* $G(V) = \min\limits_{X \in L(\mathcal{H})} G(X)$,
2. V *es una* W*-inversa de* A *en* $R(B)$,
3. V *es una solución de la ecuación normal*
$$A^*W(AX - B) = 0.$$

Demostración. 1. $\Leftrightarrow$ 2. : Se sigue de la demostración del Teorema 6.1.9.
2. $\Leftrightarrow$ 3. : Fue probado en el Teorema 4.2.3. □

Teorema 6.1.11. *Sea* $A \in CR(\mathcal{H})$, $B \in L(\mathcal{H})$ *y* $W \in L(\mathcal{H})^+$, *tal que* $W^{1/2} \in S_p$, *para algún* p *con* $1 \leq p < \infty$. *El Problema* (6.3) *admite un mínimo si y sólo si* $R(B) \subseteq R(A) + R(A)^{\perp_W}$.

En este caso
$$\min_{X \in L(\mathcal{H})} \|AX - B\|_{p,W} = \|W^{1/2}_{/R(A)} B\|_p.$$
Más aún, $X_0 \in L(\mathcal{H})$ *satisface*
$$\|AX_0 - B\|_{p,W} = \|W^{1/2}_{/R(A)} B\|_p,$$
si y sólo si X_0 *es una* W*-inversa de* A *en* $R(B)$.

Demostración. Si $R(B) \subseteq R(A) + R(A)^{\perp_W}$, por el Teorema 6.1.9, existe $X_0 \in L(\mathcal{H})$ tal que $G(X_0) = \min\limits_{X \in L(\mathcal{H})} G(X) = B^* W_{/R(A)} B$, es decir
$$G(X_0) = B^* W_{/R(A)} B \leq G(X), \text{ para cada } X \in L(\mathcal{H}).$$
Dado que $W^{1/2} \in S_p$, por la Proposición 3.1.15, se sigue que
$$\|W^{1/2}_{/R(A)} B\|_p = \|W^{1/2}(AX_0 - B)\|_p = \|AX_0 - B\|_{p,W} \leq \|AX - B\|_{p,W},$$
para cada $X \in L(\mathcal{H})$, entonces
$$\min_{X \in L(\mathcal{H})} \|AX - B\|_{p,W} = \|AX_0 - B\|_{p,W} = \|W^{1/2}_{/R(A)} B\|_p.$$

La recíproca se prueba de manera similar a como se realizó en el Teorema 6.1.5.

Finalmente, si $X_0 \in L(\mathcal{H})$ minimiza (6.3), probamos que X_0 es una solución de la ecuación normal $A^*W(AX - B) = 0$ y por el Teorema 4.2.3, es una W-inversa de A en $R(B)$. Recíprocamente, si X_0 es una W-inversa de A en $R(B)$, entonces por la Proposición 6.1.10, X_0 minimiza (6.4), y por la Proposición 3.1.15, minimiza (6.3). □

Observación 6.1.12. 1. Sea $A \in CR(\mathcal{H})$ y $W \in L(\mathcal{H})^+$. El Problema (6.4) tiene un mínimo para todo $B \in L(\mathcal{H})$ si y sólo si el par $(W, R(A))$ es compatible. De hecho, si el par $(W, R(A))$ es compatible entonces para cada $B \in L(\mathcal{H})$, tenemos $R(B) \subseteq \mathcal{H} = R(A) + R(A)^{\perp_W}$ (ver Teorema 4.1.3), y por el Teorema 6.1.9, el Problema (6.4) tiene un mínimo para todo $B \in L(\mathcal{H})$. La recíproca se sigue del Teorema 6.1.2 tomando $B = I$.

2. Sea $A \in CR(\mathcal{H})$, $B \in L(\mathcal{H})$ con $\overline{R(B)} = \mathcal{H}$ y $W \in L(\mathcal{H})^+$. Si el Problema (6.4) tiene un mínimo, entonces el par $(W, R(A))$ es cuasi-compatible (ver Sección 4.1, para definiciones y propiedades de cuasi-compatibilidad y ver Sección 2.1.3, para generalidades de operadores no acotados).

   De hecho, si el Problema (6.4) tiene un mínimo, entonces por el Teorema 6.1.9, $R(B) \subseteq R(A) + R(A)^{\perp_W}$, por lo tanto
   $\mathcal{H} = \overline{R(B)} \subseteq \overline{R(A) + R(A)^{\perp_W}}$.
   Sea $\mathcal{N} = R(A) \cap R(A)^{\perp_W}$. Notar que

$$R(A) + R(A)^{\perp_W} = R(A) \dot{+} (R(A)^{\perp_W} \cap \mathcal{N}^{\perp})$$

   y definamos $Q = P_{R(A)//R(A)^{\perp_W} \ominus \mathcal{N}}$. Entonces Q es una proyección cerrada con dominio denso y por la Proposición 4.1.6, Q es W- simétrica y el par $(W, R(A))$ es cuasi-compatible.

El caso de la norma de operadores

En [54], se establece una fórmula de la derivada ϕ-direccional para la norma de operadores. Más precisamente, sea

$$S_\infty = \{X \in L(\mathcal{H}) \text{ compacto, tal que } \|X\|_\infty < \infty\},$$

donde $\|X\|_\infty = \underset{f \in \mathcal{H},\ \|f\|=1}{sup} \|Xf\|$, la norma usual de operadores. El siguiente teorema proporciona una fórmula para la derivada ϕ-direccional del operador $G_\infty = \|X\|_\infty$.

Teorema 6.1.13. *Sea $G_\infty : S_\infty \to \mathbb{R}^+$, $G_\infty = \|X\|_\infty$ y sea $X, Y \in S_\infty$. Entonces G_∞ tiene una derivada $\phi -$ direccional dada por*

$$D_\phi G_\infty(X, Y) = \underset{f \in \Phi,\ \|f\|=1}{max} Re\ [e^{i\phi} \langle U^* Y f, f \rangle],$$

donde Φ es el subespacio donde X alcanza su norma, $Re(z)$ es la parte real del número complejo z y $X = U|X|$, es la descomposición polar del operador X.

Demostración. Ver [54, Teorema 2.6]. □

El siguiente lema, probado en [54, Corolario 2.8], caracteriza los mínimos globales del operador $\|W^{1/2}(AX - B)\|_\infty$.

Lema 6.1.14. *Sea $A \in CR(\mathcal{H})$, $B \in L(\mathcal{H})$ y $W \in L(\mathcal{H})^+$, tal que $W^{1/2} \in S_\infty$. Entonces las siguientes condiciones son equivalentes:*

1. *$X_0 \in L(\mathcal{H})$ es un mínimo global de $\|W^{1/2}(AX - B)\|_\infty$.*
2. *$\inf\limits_{0\leq\phi<2\pi} \max\limits_{f\in\Phi,\ \|f\|=1} Re\,[e^{i\phi}\langle U^*W^{1/2}AYf, f\rangle] \geq 0$, para cada $Y \in L(\mathcal{H})$, donde $W^{1/2}(AX_0 - B) = U|W^{1/2}(AX_0 - B)|$, es la descomposición polar del operador $W^{1/2}(AX_0 - B)$, y Φ es el subespacio donde $W^{1/2}(AX_0 - B)$ alcanza su norma.*
3. *Para cada $Y \in L(\mathcal{H})$, existe un vector $f_Y \in \Phi$, tal que*
$$W^{1/2}AYf_Y \perp W^{1/2}(AX_0 - B)f_Y.$$

Proposición 6.1.15. *Sea $A \in CR(\mathcal{H})$, $B \in L(\mathcal{H})$ y $W \in L(\mathcal{H})^+$, tal que $W^{1/2} \in S_\infty$. $X_0 \in L(\mathcal{H})$ es un mínimo del Problema (6.3) para $p = \infty$, si y sólo si, existe $f \in \Phi$, $f \neq 0$ tal que*
$$A^*W(AX_0 - B)f = 0,$$
donde Φ es el subespacio donde $W^{1/2}(AX_0 - B)$ alcanza su norma.

Demostración. Sea $X_0 \in L(\mathcal{H})$ un mínimo global de $\|W^{1/2}(AX - B)\|_\infty$, por el Lema 6.1.14, tenemos que para cada $Y \in L(\mathcal{H})$ existe $f_Y \in \Phi$, tal que
$$\langle Yf_Y, A^*W(AX_0 - B)f_Y \rangle = 0. \tag{6.5}$$
Sea $Y_0 = A^*W(AX_0 - B)$, entonces volviendo a (6.5) tenemos
$$\left\langle Y_0 f_{Y_0}, Y_0 f_{Y_0} \right\rangle = 0.$$
Por lo tanto
$$A^*W(AX_0 - B)f_{Y_0} = 0,$$
y obtenemos la primer conclusión.

Recíprocamante, sea $f \in \Phi$ no nulo, tal que $A^*W(AX_0 - B)f = 0$. Entonces
$$\langle Yf, A^*W(AX_0 - B)f \rangle = 0, \text{ para cada } Y \in L(\mathcal{H}).$$
Por lo tanto, por el Lema 6.1.14, X_0 es un mínimo de $\|W^{1/2}(AX - B)\|_\infty$. □

Observación 6.1.16. Sea $A \in CR(\mathcal{H})$, $B \in L(\mathcal{H})$ y $W \in L(\mathcal{H})^+$, tal que $W^{1/2} \in S_\infty$. Si $R(B) \subseteq R(A) + R(A)^{\perp_W}$, por la Proposición 4.2.3 , existe $X_0 \in L(\mathcal{H})$, tal que $A^*W(AX_0 - B) = 0$, entonces
$$A^*W(AX_0 - B)f = 0, \text{ para cada } f \in \Phi,$$
y por la Proposición 6.1.15, X_0 es un mínimo del Problema (6.3) para $p = \infty$.

6.2 SOLUCIÓN DE NORMA MÍNIMA

En esta sección estamos interesados en caracterizar entre todos los operadores que realizan el mínimo de G, aquellos que tienen norma mínima. Dados $A \in CR(\mathcal{H})$, $B \in L(\mathcal{H})$ y $W \in L(\mathcal{H})^+$, tal que $R(B) \subseteq R(A) + R(A)^{\perp_W}$, definimos

$$B' = (A^*WA)^\dagger A^*WB \in L(\mathcal{H}).$$

Consideremos el conjunto

$$M_B = \{X \in L(\mathcal{H}) : G(X) = \min_{X \in L(\mathcal{H})} G(X) = B^* W_{/R(A)} B\},$$

entonces por la Proposición 6.1.10, el Teorema 4.2.3 y el Corolario 4.2.4, tenemos que

$$M_B = \{X \in L(\mathcal{H}) : A^*W(AX - B) = 0\} = \\ \{X = B' + P_{N(A^*WA)}Z : Z \in L(\mathcal{H})\}. \quad (6.6)$$

Sea $W_2 \in L(\mathcal{H})^+$, tal que $W_2^{1/2} \in S_q$, para algún q con $1 \leq q < \infty$. Analizamos la existencia de

$$\min_{X \in M_B} \|X\|_{q,W_2}. \quad (6.7)$$

Teorema 6.2.1. *Sea $A \in CR(\mathcal{H})$, $W \in L(\mathcal{H})^+$, $W_2 \in L(\mathcal{H})^+$, tal que $W_2^{1/2} \in S_q$, para algún q con $1 \leq q < \infty$ y sea $B \in L(\mathcal{H})$, tal que $R(B) \subseteq R(A) + R(A)^{\perp_W}$. Entonces $X_1 \in M_B$ es un mínimo del Problema 6.7, es decir,*

$$\min_{X \in M_B} \|X\|_{q,W_2} = \|X_1\|_{q,W_2},$$

si y sólo si $R(W_2 X_1) \subseteq N(WA)^\perp$.

Demostración. Observar que, si $X \in M_B$, usando (6.6), el Problema (6.7) tiene mínimo si y sólo si existe

$$\min_{Z \in L(\mathcal{H})} \|P_{N(A^*WA)}Z + B'\|_{q,W_2}. \quad (6.8)$$

Entonces, por el Teorema 6.1.11 y la Proposición 6.1.10, $Z_1 \in L(\mathcal{H})$ es un mínimo del Problema (6.8) si y sólo si, Z_1 es tal que

$$P_{N(A^*WA)} W_2 (P_{N(A^*WA)} Z_1 + B') = 0, \quad (6.9)$$

si y sólo si $X_1 = P_{N(A^*WA)} Z_1 + B' \in M_B$, es un mínimo del Problema (6.7) y

$$P_{N(A^*WA)} W_2 X_1 = 0,$$

si y sólo si $R(W_2 X_1) \subseteq N(A^*WA)^\perp = N(WA)^\perp$. □

A continuación, definiremos las WW_2-inversas de A. En el contexto de espacios de dimensión finita, dichas inversas son llamadas las inversas de mínima seminorma de cuadrados mínimos [73, 66].

Definición. Dados $A \in CR(\mathcal{H})$, $B \in L(\mathcal{H})$ y $W, W_2 \in L(\mathcal{H})^+$, $X_1 \in L(\mathcal{H})$ es una WW_2-inversa de A en $R(B)$, si X_1 es una W-inversa de A en $R(B)$ y para cada $x \in \mathcal{H}$, $X_1 x$ tiene mínima W_2-seminorma entre las W-LSS de $Az = Bx$, es decir

$$\|X_1 x\|_{W_2} \leq \|u\|_{W_2}, \text{ para cada } x \in \mathcal{H} \text{ y para cada } u \text{ } W\text{-LSS de } Az = Bx.$$

Cuando $B = I$, X_1 se llama WW_2-inversa de A, ver [20].

Proposición 6.2.2. *Sea $A \in CR(\mathcal{H})$, $B \in L(\mathcal{H})$ y $W, W_2 \in L(\mathcal{H})^+$, tal que $W_2^{1/2} \in S_q$ para algún q, con $1 \leq q < \infty$. Supongamos que $R(B) \subseteq R(A) + R(A)^{\perp_W}$, entonces X_1 realiza el mínimo del Problema (6.7), es decir, $X_1 \in M_B$ es tal que*

$$\min_{X \in M_B} \|X\|_{q,W_2} = \|X_1\|_{q,W_2},$$

si y sólo si X_1 es una WW_2-inversa de A en $R(B)$.

Demostración. Supongamos que X_1 es una WW_2-inversa de A en $R(B)$, entonces como X_1 es una W-inversa de A en $R(B)$, por la Proposición 6.1.10, tenemos que $X_1 \in M_B$. Por otra parte,

$$\|X_1 x\|_{W_2} \leq \|u\|_{W_2}, \text{ para cada } x \in \mathcal{H} \text{ y para cada } u \text{ } W\text{-LSS de } Az = Bx.$$

Sea $X \in M_B$, entonces por la Proposición 6.1.10 y el Corolario 4.2.4,

$$X = B' + L,$$

donde $R(L) \subseteq N(A^*WA)$. Sea $x \in \mathcal{H}$, entonces por el Teorema 4.2.1, Xx es una W-LSS de $Az = Bx$, para cada $X \in M_B$ y para cada $x \in \mathcal{H}$. Entonces

$$\|X_1 x\|_{W_2} \leq \|Xx\|_{W_2}, \text{ para cada } x \in \mathcal{H} \text{ y para cada } X \in M_B,$$

por lo tanto $X_1^* W_2 X_1 = \min_{X \in M_B} X^* W_2 X$. Finalmente, como $W_2^{1/2} \in S_q$ para algún q, con $1 \leq q < \infty$, usando la Proposición 3.1.15, se concluye que

$$\min_{X \in M_B} \|X\|_{q,W_2} = \|X_1\|_{q,W_2}.$$

Recíprocamente, si $X_1 \in M_B$ es tal que

$$\min_{X \in M_B} \|X\|_{q,W_2} = \|X_1\|_{q,W_2},$$

por la Proposición 6.1.10, X_1 es una W-inversa de A en $R(B)$. Por otra parte, por el Teorema 6.2.1, tenemos $R(W_2 X_1) \subseteq N(WA)^\perp$.

Sea $x \in \mathcal{H}$, entonces por el Teorema 4.2.1, cualquier W-LSS de $Az = Bx$, se puede escribir como

$$u = X_1 x + h,$$

con $h \in N(A^*WA)$. Por lo tanto

$$\|u\|^2_{W_2} = \langle\, W_2 u, u \,\rangle = \langle\, W_2(X_1 x + h), X_1 x + h \,\rangle =$$

$$= \langle\, W_2 X_1 x, X_1 x \,\rangle + \langle\, W_2 h, h \,\rangle \geq \langle\, W_2 X_1 x, X_1 x \,\rangle = \|X_1 x\|_{W_2},$$

para cada $x \in \mathcal{H}$, donde usamos $R(W_2 X_1) \subseteq N(A^*WA)^\perp$. Luego X_1 es una WW_2-inversa de A en $R(B)$. $\square$

En esta sección, estudiamos también, el siguiente problema de interés, que está relacionado con el Problema (6.7): dados $A \in CR(\mathcal{H})$, $W, W_2 \in L(\mathcal{H})^+$ y $B \in L(\mathcal{H})$, analizar la existencia de

$$\underset{X \in M_B}{min}\ X^* W_2 X, \tag{6.10}$$

donde el orden es el inducido por el cono de operadores positivos de $L(\mathcal{H})$.

Teorema 6.2.3. *Sea $A \in CR(\mathcal{H})$, $W, W_2 \in L(\mathcal{H})^+$, y $B \in L(\mathcal{H})$, tal que $R(B) \subseteq R(A) + R(A)^{\perp_W}$. Entonces $X_1 \in M_B$ es un mínimo del Problema (6.10), es decir,*

$$X_1^* W_2 X_1 = \underset{X \in M_B}{min}\ X^* W_2 X$$

si y sólo si $R(W_2 X_1) \subseteq N(WA)^\perp$.

Demostración. Observar que, si $X \in M_B$, usando (6.6), el Problema (6.10) tiene mínimo si y sólo existe

$$\underset{Z \in L(\mathcal{H})}{min} (P_{N(A^*WA)} Z + B')^* W_2 (P_{N(A^*WA)} Z + B'), \tag{6.11}$$

donde el orden es el inducido por el cono de operadores positivos de $L(\mathcal{H})$. Entonces, con argumentos similares a los usados en la prueba del Teorema 6.2.1, se concluye que X_1 es un mínimo del Problema (6.10) si y sólo si $R(W_2 X_1) \subseteq N(A^*WA)^\perp = N(WA)^\perp$. $\square$

6.2.1 *Caso $B = I$*

A continuación, caracterizaremos los operadores que minimizan el Problema (6.7) para el caso $B = I$ (o para el caso en que el operador B es sobreyectivo). En este caso, notamos $M_I = M$.

Teorema 6.2.4. *Sea $A \in CR(\mathcal{H})$ y $W, W_2 \in L(\mathcal{H})^+$, tal que $W_2^{1/2} \in S_q$ para algún q, con $1 \leq q < \infty$. Entonces las siguientes condiciones son equivalentes:*

1. *Los pares $(W, R(A))$ y $(W_2, N(WA))$ son compatibles,*

2. *el Problema (6.7) tiene mínimo, es decir, existe $X_1 \in M$ tal que*

$$\underset{X \in M}{min} \|X\|_{q,W_2} = \|X_1\|_{q,W_2},$$

3. *el operador A admite una WW_2-inversa.*

Demostración.

1. $\Rightarrow$ 2. : Dado que el par $(W, R(A))$ es compatible, por el Teorema 4.1.3 y la Proposición 4.2.3, existe $D \in L(\mathcal{H})$ una solución de la ecuación $A^*W(AX - I) = 0$ y dado que el par $(W_2, N(WA))$ es compatible, sea $T_2 \in P(W_2, N(WA))$.

Definamos
$$X_1 = (I - T_2)A^\dagger T_1,$$
con $T_1 = AD$. Entonces
$$A^*WAX_1 = A^*WAA^\dagger AD - A^*WAT_2A^\dagger T_1 = A^*WAD = A^*W,$$
donde usamos $R(T_2) = N(WA)$. Por lo tanto, $X_1 \in M$.

Finalmente, tenemos que
$$R(W_2X_1) = W_2R(X_1) \subseteq W_2R(I - T_2) \subseteq W_2[R(W_2T_2)^\perp] =$$
$$W_2[(W_2N(WA))^\perp] = W_2[W_2^{-1}(N(WA)^\perp)] \subseteq N(WA)^\perp,$$
donde usamos $R(T_2) = N(WA)$ y ya que $W_2T_2 = T_2^*W_2$ y $T_2^2 = T_2$, tenemos $R(I - T_2) \subseteq R(W_2T_2)^\perp$. Por lo tanto por el Teorema 6.2.1,
$$\min_{X \in M_B} \|X\|_{q,W_2} = \|X_1\|_{q,W_2}.$$

2. $\Rightarrow$ 3. : Se sigue de la Proposición 6.2.2

3. $\Rightarrow$ 1. : Sea X_1 una WW_2-inversa de A, entonces X_1 es una W-inversa de A y por el Teorema 4.2.3, tenemos que $\mathcal{H} = R(A) + R(A)^{\perp_W}$, luego por el Teorema 4.1.3, el par $(W, R(A))$ es compatible.

Por otro lado, si definimos $\mathcal{N} = R(A)^{\perp_W} \cap R(A)$, tenemos que $\mathcal{H} = R(A) \dot{+} (R(A)^{\perp_W} \ominus \mathcal{N})$ y la proyección oblicua $Q = P_{R(A)//R(A)^{\perp_W} \ominus \mathcal{N}}$, está bien definida.

Dado $x \in \mathcal{H}$, sea $u_x = A^\dagger Qx$, luego para cada $z \in \mathcal{H}$, $u_x + P_{N(WA)}z$, es una W-LSS de $Az = x$. De hecho, $A^*WA(u_x + P_{N(WA)}z) = A^*WAA^\dagger Qx = A^*WQx = A^*Wx$ y por el Teorema 4.2.1 obtenemos la conclusión. Por lo tanto, dado que X_1 es una WW_2-inversa de A, tenemos que
$$\|X_1x\|_{W_2} = \min_{z \in \mathcal{H}} \|u_x + P_{N(WA)}z\|_{W_2}. \tag{6.12}$$

Observar que, dado que X_1 es una W-inversa de A, por el Teorema 4.2.3, X_1 es una solución de la ecuación $A^*W(AX - I) = 0$. Por lo tanto, existe $z_x \in N(WA)$ tal que $X_1x = u_x + z_x = u_x + P_{N(WA)}z_x$. Volviendo a (6.12), tenemos que
$$\|u_x + P_{N(WA)}z_x\|_{W_2} = \min_{z \in \mathcal{H}} \|u_x + P_{N(WA)}z\|_{W_2}.$$
Por lo tanto z_x es una W_2-LSS de $P_{N(WA)}z = -u_x$.

Observar que $\{u_x : x \in \mathcal{H}\} = \{A^\dagger Qx : x \in \mathcal{H}\} = R(A^\dagger Q) = A^\dagger R(A) = N(A)^\perp$. Más aún, dado que $N(A) \subseteq N(WA)$, la ecuación $P_{N(WA)}z = -u_x$, admite una solución exacta (la cual es también una W_2-LSS) para cada $x \in N(A)$. Por lo tanto, la ecuación $P_{N(WA)}z = -u_x$ admite una W_2-LSS para cada $x \in \mathcal{H}$, entonces por el Teorema 4.2.2, el par $(W_2, N(WA))$ es compatible. □

Proposición 6.2.5. *Sea $A \in CR(\mathcal{H})$ y $W, W_2 \in L(\mathcal{H})^+$. Supongamos que el par $(W, R(A))$ es compatible, entonces si X_1 es una WW_2-inversa de A, entonces X_1 es un W-inversa de A y A es una W_2-inversa de X_1.*

Demostración. Si $X_1 \in L(\mathcal{H})$ es una WW_2-inversa de A, entonces X_1 es una $W-$inversa de A y por la Proposición 6.1.10, X_1 cumple que

$$A^*WAX_1 = A^*W,$$

es decir

$$WAX_1 = (AX_1)^*WAX_1 \geq 0,$$

por lo tanto AX_1 es W-autoadjunto. Entonces

$$WA = (AX_1)^*WA = WAX_1A,$$

entonces tenemos

$$A^*WA = A^*WAX_1A,$$

ó equivalentemente $R(I - X_1A) \subseteq N(A^*WA)$. Por lo tanto, para cada $x \in \mathcal{H}$, por el Teorema 4.2.1, se sigue que

$$u = X_1x + (I - X_1A)h, \text{ para cada } h \in \mathcal{H},$$

es una W-LSS de $Az = x$.

Dado que X_1 es una WW_2-inversa de A, tenemos que

$$\|X_1x\|_{W_2} \leq \|X_1x + (I - X_1A)h\|_{W_2}, \text{ para cada } x,\ h \in \mathcal{H},$$

ó equivalentemente

$$\|W_2^{1/2}X_1x\| \leq \|W_2^{1/2}(X_1x + (I - X_1A)h)\|, \text{ para cada } x,\ h \in \mathcal{H},$$

por lo tanto

$$\langle W_2X_1x, (I - X_1A)h \rangle = 0, \text{ para cada } x,\ h \in \mathcal{H},$$

ó

$$X_1^*W_2 = X_1^*W_2X_1A,$$

y A es una W_2-inversa de X_1. □

6.3 SOLUCIONES DE CUADRADOS MÍNIMOS CON PESO DE LA ECUACIÓN AXB-C$=0$

En lo que sigue, el símbolo T^- representa una inversa interna arbitraria de $T \in L(\mathcal{H})$, es decir $T^- : D(T^-) \subseteq \mathcal{H} \to \mathcal{H}$, con $D(T^-)$, el dominio de T^-, el cual cumple $R(T) \subseteq D(T^-)$ y

$$TT^-T = T.$$

Tener en cuenta que en general $T^- \notin L(\mathcal{H})$, (ver Sección 2.1.3, para generalidades de operadores no acotados). De hecho, dada $T \in L(\mathcal{H})$ existe una inversa interna de T, T^-, tal que $T^- \in L(\mathcal{H})$, si y sólo si T tiene rango cerrado [67]. Si además, T^- verifica

$$T^-TT^- = T^-,$$

entonces T^- se denomina una inversa generalizada de T, ver Sección 2.1.4. Más aún, existe una única inversa generalizada que además verifica

$$(TT^-)^* = TT^-$$

y

$$(T^-T)^* = T^-T,$$

la cual llamamos inversa generalizada de Moore-Penrose de T y denotamos como $T^\dagger$. Por lo tanto, $T^\dagger$ es la única inversa generalizada de T tal que

$$T^\dagger T = P_{\overline{R(T^*)}} \text{ y } TT^\dagger = P_{\overline{R(T)}}|_{R(T)\oplus R(T)^\perp}.$$

Siguiendo [7], daremos condiciones suficientes y necesarias para la existencia de solución de la ecuación $AXB - C = 0$, con $A, B, C \in L(\mathcal{H})$, y alguno de los operadores A, B, C con rango cerrado.

Teorema 6.3.1. *Sea $A, B, C \in L(\mathcal{H})$. Si $R(A), R(B)$ ó $R(C)$ es cerrado, entonces las siguientes afirmaciones son equivalentes:*

1. *La ecuación $AXB - C = 0$ admite solución,*
2. *$AA^-CB^-B - C = 0$, para toda inversa interna A^-, B^- de A y B respectivamente,*
3. *$R(C) \subseteq R(A)$ y $R(C^*) \subseteq R(B^*)$.*

Si alguna de estas condiciones ocurre, la solución general de la ecuación $AXB - C = 0$ es

$$X = A^-CB^- + L - A^-ALBB^-, \tag{6.13}$$

para $L \in L(\mathcal{H})$ arbitrario.

Demostración. 1. $\Rightarrow$ 2. Primero, notar que si T^- es una inversa interna de $T \in L(\mathcal{H})$ entonces $(TT^-)^2 = TT^-$ con $R(TT^-) = R(T)$ y $(T^-T)^2 = T^-T$ con $N(T^-T) = N(T)$. Ahora, si la ecuación $AXB - C = 0$ tiene solución, entonces $R(C) \subseteq R(A)$ y $N(B) \subseteq N(C)$. Por lo tanto, es inmediato ver que $AA^-CB^-B - C = 0$, para toda inversa interna A^-, B^- de A y B respectivamente.
2. $\Rightarrow$ 3. Trivial.
3. $\Rightarrow$ 1. Si $R(C^*) \subseteq R(B^*)$ entonces, por el Teorema de Douglas 2.1.1, existe $D \in L(\mathcal{H})$ tal que $B^*D = C^*$ y $N(D) = N(C^*)$ ó, equivalentemente, $\overline{R(D^*)} = \overline{R(C)}$. Entonces, $\overline{R(D^*)} \subseteq \overline{R(A)}$. Por lo tanto, si $R(A)$ ó $R(C)$ es cerrado, entonces $R(D^*) \subseteq R(A)$. Por lo tanto, como $R(D^*) \subseteq D(A^\dagger) = R(A) \oplus R(A)^\perp$, tenemos que $A^\dagger D^* \in L(\mathcal{H})$ y $A(A^\dagger D^*)B = P_{\overline{R(A)}}|_{R(A)\oplus R(A)^\perp} C = C$. Entonces la ecuación $AXB - C = 0$ tiene solución. Si $R(B)$ es cerrado, se puede hacer una prueba similar considerando la solución reducida de Douglas de la ecuación $AX = C$.

Si alguna de las condiciones anteriores ocurre, es fácil ver que la solución de la ecuación $AXB - C = 0$ es de la forma (6.13). $\square$

6.3.1 *Resultados de minimización en el orden de operadores*

En esta sección, estamos interesados en estudiar el siguiente problema de aproximación de operadores: dados $A, B \in CR(\mathcal{H})$, $C \in L(\mathcal{H}), W \in L(\mathcal{H})^+$ y $H : L(\mathcal{H}) \to L(\mathcal{H})$,

$$H(X) = (AXB - C)^* W (AXB - C),$$

analizamos la existencia de

$$\underset{X \in L(\mathcal{H})}{inf} \; H(X), \tag{6.14}$$

con el orden inducido por el cono de operadores positivos en $L(\mathcal{H})$.

En esta sección supondremos que $B \neq 0$, pues si $B = 0$ el problema a estudiar resulta trivial.

Los resultados presentados en esta sección fueron publicados en [17].

Fue probado en la Proposición 6.1.8, que si $A \in CR(\mathcal{H})$, $C \in L(\mathcal{H})$ y $W \in L(\mathcal{H})^+$, entonces

$$\underset{X \in L(\mathcal{H})}{inf} \; (AX - C)^* W (AX - C) = C^* W_{/R(A)} C.$$

Por lo tanto,

$$H(X) \geq C^* W_{/R(A)} C, \text{ para cada } X \in L(\mathcal{H}). \tag{6.15}$$

El siguiente resultado provee condiciones suficientes para la existencia del ínfimo del Problema (6.14).

Proposición 6.3.2. *Sea* $A, B \in CR(\mathcal{H})$, $C \in L(\mathcal{H})$ *y* $W \in L(\mathcal{H})^+$. *Si* $N(B) \subseteq N(A^*WC)$ *entonces el ínfimo del conjunto* $\{H(X) : X \in L(\mathcal{H})\}$ *existe y*

$$\inf_{X \in L(\mathcal{H})} H(X) = C^* W_{/R(A)} C.$$

Demostración. Supongamos que $N(B) \subseteq N(A^*WC)$. Luego, se puede verificar que

$$H(X) = G(X) + C^*WC - P_{N(B)^\perp} C^*WCP_{N(B)^\perp}, \tag{6.16}$$

donde $G(X) = (AXB - CP_{N(B)^\perp})^* W (AXB - CP_{N(B)^\perp})$.

El conjunto $\{G(X) : X \in L(\mathcal{H})\}$ siempre admite ínfimo. De hecho, sea $X \in L(\mathcal{H})$, escribiendo $W = W_{/R(A)} + W_{R(A)}$, se sigue que

$$\begin{aligned} G(X) &= (AXB - CP_{N(B)^\perp})^* W (AXB - CP_{N(B)^\perp}) \\ &= P_{N(B)^\perp} C^* W_{/R(A)} CP_{N(B)^\perp} + (AXB - CP_{N(B)^\perp})^* W_{R(A)} (AXB - CP_{N(B)^\perp}) \\ &\geq P_{N(B)^\perp} C^* W_{/R(A)} CP_{N(B)^\perp}, \end{aligned}$$

pues $R(A) \subseteq N(W_{/R(A)})$ (ver Corolario 5.1.7). Por lo tanto $P_{N(B)^\perp} C^* W_{/R(A)} CP_{N(B)^\perp}$es una cota inferior de $G(X)$.

Si $D \geq 0$ es otra cota inferior de $G(X)$, entonces

$$D \leq G(X), \text{ para cada } X \in L(\mathcal{H}).$$

En particular,

$$D \leq P_{N(B)^\perp} C^* E^* WECP_{N(B)^\perp},$$

donde E es cualquier proyección tal que $N(E) = R(A)$. De hecho $R((I - E)CP_{N(B)^\perp}) \subseteq R(I - E) = N(E) = R(A)$ y $R(P_{N(B)^\perp} C^*(I - E)^*) \subseteq R(P_{N(B)^\perp}) = R(B^*)$, entonces por el Teorema 6.3.1, existe $X_0 \in L(\mathcal{H})$, tal que $(I - E)CP_{N(B)^\perp} = AX_0B$, es decir, $(-E)CP_{N(B)^\perp} = AX_0B - CP_{N(B)^\perp}$. Entonces, por el Lema 6.1.7

$$D \leq inf\{P_{N(B)^\perp} C^* E^* WECP_{N(B)^\perp} : \ E^2 = E, \ N(E) = R(A)\} =$$

$$P_{N(B)^\perp} C^* W_{/R(A)} CP_{N(B)^\perp}.$$

Por lo tanto,

$$P_{N(B)^\perp} C^* W_{/R(A)} CP_{N(B)^\perp} = \inf_{X \in L(\mathcal{H})} G(X).$$

Entonces, se sigue que el ínfimo de $H(X)$ existe, más aún

$$\begin{aligned} \inf_{X \in L(\mathcal{H})} H(X) &= \inf_{X \in L(\mathcal{H})} G(X) + C^*WC - P_{N(B)^\perp} C^*WCP_{N(B)^\perp} \\ &= P_{N(B)^\perp} C^* W_{/R(A)} CP_{N(B)^\perp} + C^*WC - P_{N(B)^\perp} C^*WCP_{N(B)^\perp} \\ &= C^*WC - P_{N(B)^\perp} C^* W_{R(A)} CP_{N(B)^\perp}. \end{aligned}$$

Pero dado que $N(B) \subseteq N(A^*WC)$ y que $N(W_{R(A)}) = N(A^*W)$ (ver Observación 5.1.11), tenemos que $W_{R(A)}CP_{N(B)} = 0$.

Por lo tanto

$$\begin{aligned} \underset{X\in L(\mathcal{H})}{inf} \; H(X) &= C^*WC - P_{N(B)^\perp}C^*W_{R(A)}CP_{N(B)^\perp} \\ &= C^*WC - C^*W_{R(A)}C = C^*W_{/R(A)}C. \end{aligned}$$

□

Ahora damos un ejemplo donde el Problema (6.14) admite un ínfimo, pero $N(B) \not\subseteq N(A^*WC)$, mostrando que la condición en la Proposición 6.3.2 no es necesaria.

Ejemplo 1. Sea $\mathcal{H} = \mathbb{C}^2$, $W = I$, $A = C = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ y $B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. Observar que $N(B) \not\subseteq N(A^*WC)$.

Sea $X = \begin{bmatrix} x & y \\ z & w \end{bmatrix} \in \mathbb{C}^{2\times 2}$, donde $x, y, z, w \in \mathbb{C}$. Entonces

$AXB - C = \begin{bmatrix} -1 & x \\ 0 & 0 \end{bmatrix}$, y $H(X) = (AXB - C)^*(AXB - C) = \begin{bmatrix} 1 & -x \\ -\overline{x} & |x|^2 \end{bmatrix}$.

Sea $u, v \in \mathbb{C}$ entonces se puede verificar que

$$\langle (AXB - C)^*(AXB - C)(u, v), (u, v)\rangle = |u - xv|^2.$$

Dado que para cualquier $u \in \mathbb{C}$ y para cualquier $v \in \mathbb{C} \setminus \{0\}$ existe $x \in \mathbb{C}$ tal que $u - xv = 0$, se sigue que

$$\underset{X\in L(\mathcal{H})}{inf} \; H(X) = 0.$$

Ahora damos condiciones equivalentes para la existencia del mínimo del Problema (6.14).

Teorema 6.3.3. *Sea $A, B \in CR(\mathcal{H})$, $C \in L(\mathcal{H})$ y $W \in L(\mathcal{H})^+$. Entonces las siguientes son equivalentes:*

1. *El Problema (6.14) alcanza un mínimo, es decir, existe $X_0 \in L(\mathcal{H})$ tal que*

$$\underset{X\in L(\mathcal{H})}{min} \; H(X) = H(X_0), \tag{6.17}$$

2. *$R(C) \subseteq R(A) + R(A)^{\perp_W}$ y $N(B) \subseteq N(A^*WC)$,*

3. *la ecuación normal*

$$A^*W(AXB - C) = 0, \tag{6.18}$$

admite una solución.

Si alguna de estas condiciones ocurre, entonces

$$\min_{X \in L(\mathcal{H})} H(X) = C^* W_{/R(A)} C.$$

Más aún, el operador $X_0 \in L(\mathcal{H})$ satisface

$$\min_{X \in L(\mathcal{H})} H(X) = H(X_0),$$

si y sólo si $X_0 B$ es una W-inversa de A en $R(C)$.

Demostración. 1. $\Rightarrow$ 2. Supongamos que $H(X)$ tiene un elemento mínimo. Sea $X_0 \in L(\mathcal{H})$ tal que

$$H(X_0) \leq H(X), \text{ para cada } X \in L(\mathcal{H}),$$

ó equivalentemente

$$\|(AX_0B - C)x\|_W \leq \|(AXB - C)x\|_W, \text{ para cada } X \in L(\mathcal{H}) \text{ y cada } x \in \mathcal{H}.$$

Si $x \notin N(B)$ entonces $y = Bx \neq 0$, y dado $z \in \mathcal{H}$, existe $X \in L(\mathcal{H})$ tal que $z = Xy$. Por lo tanto

$$\|AX_0Bx - Cx\|_W \leq \|Az - Cx\|_W, \text{ para cada } z \in \mathcal{H}.$$

Entonces X_0Bx es una W-LSS de $Az = Cx$, y por el Teorema 4.2.1

$$C(\mathcal{H} \setminus N(B)) \subseteq R(A) + R(A)^{\perp_W}.$$

Observar que dado que $\mathcal{H} \setminus N(B)$ es un conjunto abierto no vacío, y que

$$\mathcal{H} \setminus N(B) \subseteq C^{-1}(R(A) + R(A)^{\perp_W}),$$

el subespacio $C^{-1}(R(A) + R(A)^{\perp_W})$ tiene interior no vacío , por lo tanto por la Proposición A.1.1,

$$\mathcal{H} = C^{-1}(R(A) + R(A)^{\perp_W}),$$

entonces,

$$R(C) \subseteq R(A) + R(A)^{\perp_W}.$$

Observar también, que el interior de $N(B) \subsetneq \mathcal{H}$ es vacío entonces, por la Proposición A.1.1, el conjunto $\mathcal{H} \setminus N(B)$ es un subconjunto denso de $\mathcal{H}$. Por lo tanto dado $y \in R(C)$, existe $x \in \mathcal{H}$ tal que $y = Cx$, y existe una sucesión $\{x_n\}_{n \geq 1} \subset \mathcal{H} \setminus N(B)$ tal que $\lim_{n \to \infty} x_n = x$. Entonces

$$\|AX_0Bx_n - Cx_n\|_W \leq \|Az - Cx_n\|_W, \text{ para cada } z \in \mathcal{H}, \text{ y para cada } n \in \mathbb{N},$$

y tomando límite en ambos lados de la desigualdad, obtenemos

$$\|AX_0Bx - Cx\|_W \leq \|Az - Cx\|_W, \text{ para cada } x, z \in \mathcal{H}.$$

Por lo tanto, $G = X_0 B$ es una W-inversa de A en $R(C)$ tal que $GP_{N(B)} = X_0 BP_{N(B)} = 0$. Entonces por el Teorema 4.2.3, $A^*WC = A^*WAG$ y multiplicando por $P_{N(B)}$ obtenemos

$$A^*WCP_{N(B)} = A^*WAGP_{N(B)} = 0,$$

y entonces $N(B) \subseteq N(A^*WC)$.
2. $\Rightarrow$ 3. Si $R(C) \subseteq R(A) + R(A)^{\perp_W}$, por el Teorema 4.2.3, existe $X_0 \in L(\mathcal{H})$ una solución de la ecuación normal

$$A^*W(AX_0 - C) = 0. \tag{6.19}$$

Dado que $N(B) \subseteq N(A^*WC)$, tenemos que

$$A^*WC = A^*WCP_{N(B)} + A^*WCP_{N(B)^\perp} = A^*WCP_{N(B)^\perp},$$

entonces multiplicando (6.19) por $P_{N(B)^\perp}$ se sigue que

$$A^*W(AX_0P_{N(B)^\perp} - CP_{N(B)^\perp}) = A^*W(A(X_0B^\dagger)B - C) = 0,$$

y entonces la ecuación (6.18), admite una solución.
3. $\Rightarrow$ 1. Sea X_0 una solución de la ecuación normal (6.18), entonces por el Teorema 4.2.3, $G_0 = X_0B$ es una W-inversa de A en $R(C)$, entonces tenemos

$$\|AG_0x - Cx\|_W \leq \|Az - Cx\|_W, \text{ para cada } x, z \in \mathcal{H}.$$

Dado $Y \in L(\mathcal{H})$, tomar $z = Yx$, entonces

$$\|AX_0Bx - Cx\|_W \leq \|AYx - Cx\|_W, \text{ para cada } Y \in L(\mathcal{H}), \text{ y cada } x \in \mathcal{H}.$$

En particular, si $Y = XB$, entonces

$$\|(AX_0B - C)x\|_W \leq \|(AXB - C)x\|_W, \text{ para cada } X \in L(\mathcal{H}), \text{ y cada } x \in \mathcal{H}.$$

Y entonces

$$H(X_0) \leq H(X), \text{ para cada } X \in L(\mathcal{H}).$$

Finalmente, X_0 es el mínimo del Problema (6.14), si y sólo si X_0 es una solución de la ecuación (6.18), si y sólo si X_0B es una W-inversa de A en $R(C)$ (ver Teorema 4.2.3). Por lo tanto, en este caso

$$H(X_0) = (AX_0B - C)^*W(AX_0B - C) = C^*W_{/R(A)}C,$$

donde usamos la Proposición 6.1.10. Entonces

$$\underset{X \in L(\mathcal{H})}{min} H(X) = H(X_0) = C^*W_{/R(A)}C.$$

$\square$

Corolario 6.3.4. *Sea $A, B \in CR(\mathcal{H})$, $C \in L(\mathcal{H})$ y $W \in L(\mathcal{H})^+$. Si $R(C) \subseteq R(A) + R(A)^{\perp_W}$ y $N(B) \subseteq N(A^*WC)$, entonces las soluciones del Problema (6.17) (ó de la ecuación (6.18)) son*

$$(A^*WA)^\dagger A^*WCB^\dagger + L - (A^*WA)^\dagger A^*WALBB^\dagger,$$

para $L \in L(\mathcal{H})$ arbitrario.

Demostración. Dado que $R(C) \subseteq R(A) + R(A)^{\perp_W}$ y $N(B) \subseteq N(A^*WC)$, por el Teorema 6.3.3, el Problema 6.17 (ó la ecuación (6.18)) admite una solución. Entonces, por el Teorema 6.3.1, obtenemos la conclusión. □

6.3.2 *Resultados de minimización en S_p*

En esta sección estudiamos el problema de aproximación siguiente: dados $A, B \in CR(\mathcal{H})$, $C \in L(\mathcal{H})$ y $W \in L(\mathcal{H})^+$ tal que $W^{1/2} \in S_p$ para algún p con $1 \leq p < \infty$, analizar la existencia de

$$\underset{X \in L(\mathcal{H})}{min} \|AXB - C\|_{p,W}, \tag{6.20}$$

donde $\|X\|_{p,W} = \|W^{1/2}X\|_p$.

Observar que, de la ecuación (6.15) y de la Proposición 3.1.15, se sigue que

$$\underset{X \in L(\mathcal{H})}{inf} \|AXB - C\|_{p,W} \geq \|W^{1/2}_{/R(A)}C\|_p.$$

La próxima proposición da condiciones suficientes para la existencia del mínimo del Problema (6.20).

Proposición 6.3.5. *Sea $A, B \in CR(\mathcal{H})$, $C \in L(\mathcal{H})$ y $W \in L(\mathcal{H})^+$, tal que $W^{1/2} \in S_p$, para algún p con $1 \leq p < \infty$. Si*

$$N(B) \subseteq N(A^*WC) \text{ y } R(C) \subseteq R(A) + R(A)^{\perp_W},$$

entonces existe $X_0 \in L(\mathcal{H})$ tal que

$$\underset{X \in L(\mathcal{H})}{min} \|AXB - C\|_{p,W} = \|AX_0B - C\|_{p,W} = \|W^{1/2}_{/R(A)}C\|_{p,W}.$$

Demostración. Si $N(B) \subseteq N(A^*WC)$ y $R(C) \subseteq R(A) + R(A)^{\perp_W}$, por el Teorema 6.3.3, existe $X_0 \in L(\mathcal{H})$ tal que $H(X_0) = \underset{X \in L(\mathcal{H})}{min} H(X) = C^*W_{/R(A)}C$, es decir

$$H(X_0) = C^*W_{/R(A)}C \leq H(X), \text{ para cada } X \in L(\mathcal{H}).$$

Dado que $W^{1/2} \in S_p$, por la Proposición 3.1.15 se sigue que

$$\|W^{1/2}_{/R(A)}C\|_p = \|W^{1/2}(AX_0B - C)\|_p =$$

$$= \|AX_0B - C\|_{p,W} \leq \|AXB - C\|_{p,W}, \text{ para cada } X \in L(\mathcal{H}),$$

entonces

$$\min_{X \in L(\mathcal{H})} \|AXB - C\|_{p,W} = \|AX_0B - C\|_{p,W} = \|W^{1/2}_{/R(A)}C\|_p.$$

□

Observación 6.3.6. Sea $|||\cdot|||$ una norma unitariamente invariante en un ideal no nulo $\mathcal{J}$ de $L(\mathcal{H})$. Dado $W \in L(\mathcal{H})^+$ tal que $W^{1/2} \in \mathcal{J}$. Si $N(B) \subseteq N(A^*WC)$ y $R(C) \subseteq R(A) + R(A)^{\perp_W}$, por la Proposición 3.1.15 y el Teorema 6.3.3, existe $X_0 \in L(\mathcal{H})$ tal que

$$\min_{X \in L(\mathcal{H})} \|AXB - C\|_W = \|AX_0B - C\|_W = |||W^{1/2}_{/R(A)}C|||,$$

es decir, en este caso, el Problema (6.20) tiene mínimo para cualquier norma unitariamente invariante.

Lema 6.3.7. *Sea $A, B \in CR(\mathcal{H})$, $C \in L(\mathcal{H})$ y $W \in L(\mathcal{H})^+$, tal que $W^{1/2} \in S_p$ para algún p con $1 < p < \infty$ y considerar $F_p(X) = \|AXB - C\|^p_{p,W}$. Entonces, $X_0 \in L(\mathcal{H})$ es un mínimo global de F_p si y sólo si $X_0 \in L(\mathcal{H})$ es una solución de*

$$B|W^{1/2}(AXB - C)|^{p-1}U^*W^{1/2}A = 0, \tag{6.21}$$

donde $W^{1/2}(AXB - C) = U|W^{1/2}(AXB - C)|$ es la descomposición polar del operador $W^{1/2}(AXB - C)$, con U una isometría parcial con $N(U) = N(W^{1/2}(AXB - C))$.

Demostración. Observar que en la ecuación (6.21), U depende de X.

Supongamos que $X_0 \in L(\mathcal{H})$ es un mínimo de F_p. Sea $W^{1/2}(AX_0B - C) = U|W^{1/2}(AX_0B - C)|$ la descomposición polar del operador $W^{1/2}(AX_0B - C)$, con U una isometría parcial con $N(U) = N(W^{1/2}(AX_0B - C))$. Por el Teorema 3.2.7, F_p tiene una derivada $\phi-$direccional para todo $\phi \in [0, 2\pi)$. Entonces es fácil verificar que, para cada X, $Y \in L(\mathcal{H})$ y $\phi \in [0, 2\pi)$,

$$D_\phi F_p(X, Y) = D_\phi G_p(W^{1/2}(AXB - C), W^{1/2}AYB),$$

donde $G_p(X) = \|X\|^p_p$. Por el Lema 3.2.4 y el Corolario 3.2.8, vale que, para cada $\phi \in [0, 2\pi)$

$$0 \leq D_\phi F_p(X_0, Y) = p \; Re \; [e^{i\phi} tr(|W^{1/2}(AX_0B - C)|^{p-1}U^*W^{1/2}AYB)],$$

para cada $Y \in L(\mathcal{H})$. Considerando ϕ e Y adecuados, se sigue que

$$B|W^{1/2}(AX_0B - C)|^{p-1}U^*W^{1/2}A = 0.$$

Recíprocamente, supongamos que $X_0 \in L(\mathcal{H})$ es una solución de (6.21), entonces para cada $\phi \in [0, 2\pi)$ e $Y \in L(\mathcal{H})$ tenemos

$$D_\phi F_p(X_0, Y) = 0.$$

Si $F_p(X_0) = 0$ entonces X_0 es un mínimo de F_p. Supongamos que $F_p(X_0) \neq 0$ y sea $f_p(X) = F_p(X)^{\frac{1}{p}}$, entonces

$$D_\phi f_p(X_0, Y) = 0, \text{ para cada } Y \in L(\mathcal{H}),$$

donde usamos la regla de la cadena probada en la Proposición 3.2.9.

Sea $g_p(X) = \|X\|_p$, entonces es fácil ver que para cada $Y \in L(\mathcal{H})$, tenemos

$$0 = D_\phi f_p(X_0, e^{i(\pi-\phi)}(-Y + X_0))$$

$$= D_\phi g_p(W^{1/2}(AX_0B - C), W^{1/2}Ae^{i(\pi-\phi)}(-Y + X_0)B + e^{i(\pi-\phi)}W^{1/2}C - e^{i(\pi-\phi)}W^{1/2}C)$$
$$= D_\phi g_p(W^{1/2}(AX_0B - C), -e^{i(\pi-\phi)}W^{1/2}(AYB - C) + e^{i(\pi-\phi)}W^{1/2}(AX_0B - C)).$$

Por otra parte, observar que por la prueba del Lema 3.2.4, para cada $X \in L(\mathcal{H})$, tenemos que

$$D_\phi g_p(X, e^{i(\pi-\phi)}X) = -\|X\|_p.$$

Entonces

$$\begin{aligned}
\|W^{1/2}(AX_0B - C)\|_p &= -D_\phi g_p(W^{1/2}(AX_0B - C), e^{i(\pi-\phi)}W^{1/2}(AX_0B - C)) + \\
&+ D_\phi g_p(W^{1/2}(AX_0B - C), -e^{i(\pi-\phi)}W^{1/2}(AYB - C) + e^{i(\pi-\phi)}W^{1/2}(AX_0B - C)) \\
&\leq -D_\phi g_p(W^{1/2}(AX_0B - C), e^{i(\pi-\phi)}W^{1/2}(AX_0B - C)) + \\
&+ D_\phi g_p(W^{1/2}(AX_0B - C), e^{i(\pi-\phi)}W^{1/2}(AX_0B - C)) \\
&+ D_\phi g_p(W^{1/2}(AX_0B - C), -e^{i(\pi-\phi)}W^{1/2}(AYB - C)) \\
&= D_\phi g_p(W^{1/2}(AX_0B - C), -e^{i(\pi-\phi)}W^{1/2}(AYB - C)) \\
&\leq \| - e^{i(\pi-\phi)}W^{1/2}(AYB - C)\|_p = \|W^{1/2}(AYB - C)\|_p, \text{ para cada } Y \in L(\mathcal{H}),
\end{aligned}$$

donde usamos propiedades de $D_\phi g_p$, probadas en la Proposición 3.2.3. Entonces X_0 is un mínimo global de f_p, ó equivalentemente, X_0 es un mínimo global de F_p. □

La siguiente observación es una forma alternativa de expresar el lema anterior.

Observación 6.3.8. Sea $A, B \in CR(\mathcal{H})$, $C \in L(\mathcal{H})$ y $W \in L(\mathcal{H})^+$, tal que $W^{1/2} \in S_p$, para algún p con $1 < p < \infty$. Siguiendo los mismos argumentos que en [15, Teorema 1], se puede probar que existe $X_0 \in L(\mathcal{H})$ tal que

$$\min_{X \in L(\mathcal{H})} \|AXB - C\|_{p,W} = \|AX_0B - C\|_{p,W},$$

si y sólo si X_0 es solución de la siguiente ecuación

$$W^{1/2}AX_0B = W^{1/2}C + Q_0, \tag{6.22}$$

donde $Q_0 = U_0 Q_1^{\frac{1}{p-1}}$, con $Q_1 = |L_1 - (A^*W^{1/2})^-(A^*W^{1/2})L_1B^*(B^*)^-|$ y

$$L_1 - (A^*W^{1/2})^-(A^*W^{1/2})L_1B^*(B^*)^- = U_0Q_1,$$

la descomposición polar del operador $L_1 - (A^*W^{1/2})^-(A^*W^{1/2})L_1B^*(B^*)^-$, donde $(B^*)^- \in L(\mathcal{H})$ es una inversa interna de B^* y $(A^*W^{1/2})^- : D((A^*W^{1/2})^-) \subseteq \mathcal{H} \to \mathcal{H}$, con $R(A^*W^{1/2}) \subseteq D((A^*W^{1/2})^-)$ es una inversa interna de $A^*W^{1/2}$.

Suponiendo que $N(B) \subseteq N(A^*WC)$, es posible probar la recíproca de la Proposición 6.3.5.

Teorema 6.3.9. *Sea $A, B \in CR(\mathcal{H})$, $C \in L(\mathcal{H})$ y $W \in L(\mathcal{H})^+$, tal que $W^{1/2} \in S_p$ para algún p con $1 \leq p < \infty$ y $N(B) \subseteq N(A^*WC)$. Entonces las siguientes son equivalentes:*

1. *Existe $X_0 \in L(\mathcal{H})$ tal que*

$$\min_{X \in L(\mathcal{H})} \|AXB - C\|_{p,W} = \|AX_0B - C\|_{p,W},$$

2. *la ecuación normal*

$$A^*W(AXB - C) = 0, \tag{6.23}$$

   admite una solución.

3. $R(C) \subseteq R(A) + R(A)^{\perp_W}$,

4. *existe $X_0 \in L(\mathcal{H})$ tal que*

$$\min_{X \in L(\mathcal{H})} (AXB - C)^*W(AXB - C) = (AX_0B - C)^*W(AX_0B - C).$$

En este caso,

$$\min_{X \in L(\mathcal{H})} \|AXB - C\|_{p,W} = \|W^{1/2}_{/R(A)}C\|_p.$$

Más aún, $X_0 \in L(\mathcal{H})$ satisface

$$\|AX_0B - C\|_{p,W} = \|W^{1/2}_{/R(A)}C\|_p,$$

si y sólo si X_0 es como en el Corolario 6.3.4.

Demostración. 1. $\Rightarrow$ 2. Para $1 \leq p < \infty$, considerar $F_p : S_p \to \mathbb{R}^+$,

$$F_p(X) = \|W^{1/2}(AXB - C)\|_p^p.$$

Por el Teorema 3.2.7, F_p tiene una derivada $\phi-$direccional para todo $\phi \in [0, 2\pi)$. Entonces es fácil verificar que, para cada $X, Y \in L(\mathcal{H})$ y $\phi \in [0, 2\pi)$,

$$D_\phi F_p(X, Y) = D_\phi G_p(W^{1/2}(AXB - C), W^{1/2}AYB),$$

donde $G_p(X) = \|X\|_p^p$.

Supongamos que existe $X_0 \in L(\mathcal{H})$, un mínimo global de $\|AXB - C\|_{p,W}$. Entonces X_0 es un mínimo de F_p y, por el Lema 3.2.4, tenemos

$$\inf_{0 \leq \phi < 2\pi} (D_\phi F_p(X_0, Y)) \geq 0, \text{ para cada } Y \in L(\mathcal{H}).$$

Sea $W^{1/2}(AX_0B - C) = U|W^{1/2}(AX_0B - C)|$ la descomposición polar del operador $W^{1/2}(AX_0B - C)$, con U una isometría parcial con $N(U) = N(W^{1/2}(AX_0B - C))$, $P = P_{N(W^{1/2}(AX_0B - C))}$ y $Q = P_{N((W^{1/2}(AX_0B - C))^*)}$.

Si $p = 1$, por el Corolario 3.2.14, vale para cada $\phi \in [0, 2\pi)$

$$0 \leq D_\phi F_1(X_0, Y) = Re\,[e^{i\phi} tr(U^* W^{1/2} AYB)] + \|QW^{1/2}AYBP\|_1, \text{ para cada } Y \in L(\mathcal{H}).$$

Considerando un ϕ adecuado para cada $Y \in L(\mathcal{H})$, obtenemos

$$|tr(U^* W^{1/2} AYB)| \leq \|QW^{1/2}AYBP\|_1, \text{ para cada } Y \in L(\mathcal{H}),$$

ó equivalentemente

$$|tr(U^* W^{1/2} AZ)| \leq \|QW^{1/2}AZP\|_1, \text{ para cada } Z \in L(\mathcal{H}), \text{ con } N(B) \subseteq N(Z).$$

Observar que $R(Q) = N(U^*)$ y $R(P) = N(U)$, por lo tanto $U^*Q = PU^* = 0$. También, observar que dado que $N(B) \subseteq N(A^*WC)$, tenemos que

$$N(B) \subseteq N(A^*W(AX_0B - C)(I - P)).$$

Sea $Y \in L(\mathcal{H})$ entonces

$$|tr(U^* W^{1/2} AYA^*W(AX_0B - C))| = |tr((I - P)U^* W^{1/2} AYA^*W(AX_0B - C))| =$$
$$= |tr(U^* W^{1/2} AYA^*W(AX_0B - C)(I - P))| \leq$$
$$\|QW^{1/2}AYA^*W(AX_0B - C)(I - P)P\|_1 = 0,$$

donde usamos $N(B) \subseteq N(YA^*W(AX_0B - C)(I - P))$. Entonces

$$tr(U^* W^{1/2} AYA^*W(AX_0B - C)) = tr(A^*W(AX_0B - C)U^* W^{1/2} AY) = 0,$$

para cada $Y \in L(\mathcal{H})$. Por lo tanto

$$A^*W(AX_0B - C)U^* W^{1/2} A = A^* W^{1/2} U|W^{1/2}(AX_0B - C)|U^* W^{1/2} A = 0.$$

Entonces
$$|W^{1/2}(AX_0B-C)|^{1/2}U^*W^{1/2}A=0,$$
ó
$$|W^{1/2}(AX_0B-C)|U^*W^{1/2}A=(AX_0B-C)^*WA=0,$$
ó equivalentemente
$$A^*W(AX_0B-C)=0.$$

Si $1<p<\infty$, por el Lema 6.3.7,
$$B|W^{1/2}(AX_0B-C)|^{p-1}U^*W^{1/2}A=0.$$
Observar que, dado que $N(B)\subseteq N(A^*WC)$, tenemos que
$$N(B)\subseteq N(A^*W(AX_0B-C))=N(A^*W^{1/2}U|W^{1/2}(AX_0B-C)|). \quad (6.24)$$
Por otra parte, tenemos
$$R(|W^{1/2}(AX_0B-C)|^{p-1}U^*W^{1/2}A)\subseteq N(B),$$
y de (6.24) tenemos que
$$A^*W^{1/2}U|W^{1/2}(AX_0B-C)||W^{1/2}(AX_0B-C)|^{p-1}U^*W^{1/2}A=$$
$$A^*W^{1/2}U|W^{1/2}(AX_0B-C)|^{p}U^*W^{1/2}A=0.$$
Entonces
$$A^*W^{1/2}U|W^{1/2}(AX_0B-C)|^{p}U^*W^{1/2}A=$$
$$=A^*W^{1/2}U|W^{1/2}(AX_0B-C)|^{p/2}|W^{1/2}(AX_0B-C)|^{p/2}U^*W^{1/2}A=0.$$
Por lo tanto
$$|W^{1/2}(AX_0B-C)|^{p/2}U^*W^{1/2}A=0,$$
y dado que, por la Proposición A.2.3, $N(|W^{1/2}(AX_0B-C)|^r)=N(|W^{1/2}(AX_0B-C)|^s)$, para $s,t>0$, tenemos que
$$|W^{1/2}(AX_0B-C)|U^*W^{1/2}A=(AX_0B-C)^*WA=0,$$
ó equivalentemente
$$A^*W(AX_0B-C)=0.$$
2. $\Rightarrow$ 3. Ver Teorema 4.2.3.
3. $\Rightarrow$ 4. Se sigue del Teorema 6.3.3.
4. $\Rightarrow$ 1. Ver la prueba de la Proposición 6.3.5.

Finalmente, $X_0\in L(\mathcal{H})$ es tal que $\min\limits_{X\in L(\mathcal{H})}\|AXB-C\|_{p,W}=\|AX_0B-C\|_{p,W}$, si y sólo si X_0 es una solución de la ecuación normal (6.23), y entonces X_0 es como en el Corolario 6.3.4 y por el Teorema 6.3.3 y la Proposición 3.1.15
$$\min_{X\in L(\mathcal{H})}\|AXB-C\|_{p,W}=\|AX_0B-C\|_{p,W}=\|W^{1/2}_{/R(A)}C\|_p.$$

□

Observar que la ecuación $A^*W(AXB-C)=0$ admite una solución, si y sólo si la ecuación $A^*W(AXB-C)B^*=0$ admite una solución y $N(B)\subseteq N(A^*WC)$. Entonces cuando la ecuación $A^*W(AXB-C)=0$ admite una solución, el conjunto de soluciónes de $A^*W(AXB-C)=0$ y de la ecuación $A^*W(AXB-C)B^*=0$ coinciden.

Observar además, que si $N(B)\subseteq N(A^*WC)$ entonces $R(C)\subseteq R(A)+R(A)^{\perp_W}$ si y sólo si $R(CB^*)\subseteq R(A)+R(A)^{\perp_W}$.

Cuando $p=2$, es posible caracterizar la existencia del mínimo del Problema (6.20) sin hipótesis adicionales.

Teorema 6.3.10. *Sea $A,B\in CR(\mathcal{H})$, $C\in L(\mathcal{H})$ y $W\in L(\mathcal{H})^+$, tal que $W^{1/2}\in S_2$. Entonces las siguientes son equivalentes:*

1. *Existe el mínimo del Problema (6.20) para $p=2$, es decir, existe $X_0\in L(\mathcal{H})$ tal que*
$$\min_{X\in L(\mathcal{H})}\|AXB-C\|_{2,W}=\|AX_0B-C\|_{2,W},$$
2. *la ecuación normal*
$$A^*W(AXB-C)B^*=0, \tag{6.25}$$
admite solución.
3. $R(CB^*)\subseteq R(A)+R(A)^{\perp_W}$.

En este caso,
$$\min_{X\in L(\mathcal{H})}\|AXB-C\|_{2,W}=\|W^{1/2}_{/R(A)}CP_{N(B)^\perp}+W^{1/2}CP_{N(B)}\|_2.$$
Más aún, $X_0\in L(\mathcal{H})$ satisface
$$\|AX_0B-C\|_{2,W}=\|W^{1/2}_{/R(A)}CP_{N(B)^\perp}+W^{1/2}CP_{N(B)}\|_2,$$
si y sólo si $X_0=(A^*WA)^\dagger A^*WCP_{N(B)^\perp}B^\dagger+L-(A^*WA)^\dagger A^*WALBB^\dagger$, para $L\in L(\mathcal{H})$ arbitrario.

Demostración. 1. $\Leftrightarrow$ 2. Se sigue del Lema 6.3.7.
2. $\Leftrightarrow$ 3. $R(CP_{N(B)^\perp})=R(CB^*)\subseteq R(A)+R(A)^{\perp_W}$ y $N(B)\subseteq N(A^*WCP_{N(B)^\perp})$ si y sólo si (por el Teorema 6.3.3), existe solución de la ecuación
$$A^*W(AXB-CP_{N(B)^\perp})=0, \tag{6.26}$$
si y sólo si, existe solución de la ecuación
$$A^*W(AXB-C)B^*=0.$$

Finalmente, $X_0 \in L(\mathcal{H})$ es el mínimo del Problema (6.20) para $p = 2$, si y sólo si X_0 es solución de la ecuación normal (6.25) (ó equivalentemente es solución de la ecuación (6.26)). Por el Teorema 6.3.3 y la Proposición 3.1.15, tenemos que

$$\underset{X \in L(\mathcal{H})}{min} \|AXB - CP_{N(B)^\perp}\|_{2,W} = \|AX_0B - CP_{N(B)^\perp}\|_{2,W} = \|W^{1/2}_{/R(A)} CP_{N(B)^\perp}\|_2.$$

Entonces, como

$$\|AXB - C\|_{2,W} = \|W^{1/2}(AXB - CP_{N(B)^\perp} - CP_{N(B)})\|_2,$$

y $W^{1/2}CP_{N(B)}(W^{1/2}(AXB - CP_{N(B)^\perp}))^* = 0$, tenemos que

$$\|AXB - C\|^2_{2,W} = \|AXB - CP_{N(B)^\perp}\|^2_{2,W} + \|CP_{N(B)}\|^2_{2,W}.$$

Entonces

$$\underset{X \in L(\mathcal{H})}{min} \|AXB - C\|^2_{2,W} = \underset{X \in L(\mathcal{H})}{min} \|AXB - CP_{N(B)^\perp}\|^2_{2,W} + \|CP_{N(B)}\|^2_{2,W} =$$

$$= \|W^{1/2}_{/R(A)} CP_{N(B)^\perp}\|^2_2 + \|W^{1/2}CP_{N(B)}\|^2_2 = \|W^{1/2}_{/R(A)} CP_{N(B)^\perp} + W^{1/2}CP_{N(B)}\|^2_2,$$

donde usamos que $W^{1/2}CP_{N(B)}(W^{1/2}_{/R(A)} CP_{N(B)^\perp})^* = 0$. Por lo tanto, tenemos que

$$\underset{X \in L(\mathcal{H})}{min} \|AXB - C\|_{2,W} = \|W^{1/2}_{/R(A)} CP_{N(B)^\perp} + W^{1/2}CP_{N(B)}\|_2.$$

Finalmente, $X_0 \in L(\mathcal{H})$ satisface

$$\|AX_0B - C\|_{2,W} = \|W^{1/2}_{/R(A)} CP_{N(B)^\perp} + W^{1/2}CP_{N(B)}\|_2,$$

si y sólo si es solución de la ecuación normal (6.25) (ó equivalentemente es solución de la ecuación (6.26)), entonces por el Teorema 6.3.1, $X_0 = (A^*WA)^\dagger A^*WCP_{N(B)^\perp}B^\dagger + L - (A^*WA)^\dagger A^*WALBB^\dagger$, para $L \in L(\mathcal{H})$ arbitrario.

□

La existencia de solución del Problema (6.17) implica la existencia de solución del Problema (6.20), pero el Ejemplo 2 muestra que la recíproca no es verdad. Notar que en este caso $N(B) \not\subseteq N(A^*WC)$, entonces el Problema (6.17) no admite mínimo. Este ejemplo, muestra también que, en general, para $1 < p < \infty$ un mínimo global de $F_p : S_p \to \mathbb{R}$,

$$F_p(X) = \|AXB - C\|^p_{p,W}$$

no necesariamente es una solución de la ecuación normal

$$A^*W(AXB - C)B^* = 0,$$

lo que contradice [63, Teorema 4.1], donde se afirma que $X_0 \in L(\mathcal{H})$ es un mínimo de F_p si y sólo si X_0 es una solución de la ecuación normal $A^*W(AXB - C)B^* = 0$.

Ejemplo 2. Sea $\mathcal{H} = \mathbb{C}^2$, $W = I$, la matriz identidad, $A = \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix}$, $B = \begin{pmatrix} a^2 & -1 \\ a^2 & -1 \end{pmatrix}$ y $C = \begin{pmatrix} 1 & 0 \\ 0 & a^{\frac{2}{p-1}} \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$, con $a, p > 1$.

Sea $X_0 = \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}$, entonces es fácil verificar que $AX_0B = 0$, por lo tanto

$$B|AX_0B - C|^{p-1}U^*A = B|C|^{p-1}U^*A = B \begin{pmatrix} -1 & 0 \\ 0 & a^2 \end{pmatrix} A = 0,$$

es decir, en virtud del Lema 6.3.7, X_0 es un mínimo global de F_p.

Por otra parte,

$$B(AX_0B - C)^*A = -BC^*A = -B \begin{pmatrix} -1 & 0 \\ 0 & a^{\frac{2}{p-1}} \end{pmatrix} A = (a^2 - a^{\frac{2}{p-1}}) \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \neq 0,$$

para cada $p \neq 2$. Entonces para $p \neq 2$, se sigue que X_0 es un mínimo global de F_p pero no es una solución de la ecuación normal $A^*W(AXB - C)B^* = 0$.

Por otro lado, el Ejemplo 3 muestra que en general una solución de la ecuación normal $A^*W(AXB - C)B^* = 0$, no necesariamente es un mínimo global de F_p, para $2 < p < \infty$, con $p \neq 2$. Observar que en [63, Teorema 4.1], se muestra que tampoco una solución de la ecuación normal $A^*W(AXB - C)B^* = 0$, era necesariamente un mínimo global de F_p, para $1 < p < 2$.

Ejemplo 3. Sea $\mathcal{H} = \mathbb{C}^2$, $W = I$, $A = B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ y $C = \begin{bmatrix} 1 & 1 \\ 1 & 5 \end{bmatrix}$. Sea $X_0 \in L(\mathcal{H})$ y sea $AX_0B - C = |AX_0B - C|U$, la descomposición polar de $AX_0B - C$. Se puede verificar que $X_0 = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$, es una solución de

$$0 = B|AX_0B - C|U^*A = B(AX_0B - C)^*A.$$

De hecho, si definimos $Z_0 = AX_0B - C$ entonces $Z_0^* = (AX_0B - C)^* = \begin{pmatrix} 1 & 1 \\ 1 & -3 \end{pmatrix}$, y

$$B(AX_0B - C)^*A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -3 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = 0.$$

Sea $Y_0 = (AX_0B - C)^*(AX_0B - C) = \begin{pmatrix} 2 & -2 \\ -2 & 10 \end{pmatrix}$, entonces $|Z_0| = Y_0^{1/2}$ y se puede verificar que Y_0 se puede diagonalizar de la siguiente manera

$$Y_0 = VDV^{-1},$$

con
$$V = \begin{pmatrix} 2+\sqrt{5} & 2-\sqrt{5} \\ 1 & 1 \end{pmatrix}, D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix},$$
con $\lambda_1 = 6-2\sqrt{5}, \ \ \lambda_2 = 6+2\sqrt{5}$, y
$$V^{-1} = \frac{1}{10} \begin{pmatrix} \sqrt{5} & 5-2\sqrt{5} \\ -\sqrt{5} & 5+2\sqrt{5} \end{pmatrix}.$$

Sea $p \geq 2$, entonces
$$B|AX_0B-C|^{p-1}U^*A = B|AX_0B-C|^{p-2}|AX_0B-C|U^*A =$$
$$= B|AX_0B-C|^{p-2}(AX_0B-C)^*A = B|Z_0|^{p-2}(Z_0)^*A =$$
$$= BY_0^{\frac{p-2}{2}}Z_0^*A = BVD^{\frac{p-2}{2}}V^{-1}Z_0^*A.$$

Se puede verificar que
$$BV = \begin{pmatrix} 3+\sqrt{5} & 3-\sqrt{5} \\ 3+\sqrt{5} & 3-\sqrt{5} \end{pmatrix}$$
y
$$V^{-1}Z_0^*A = \frac{1}{5} \begin{pmatrix} 3\sqrt{5}-5 & 3\sqrt{5}-5 \\ -3\sqrt{5}-5 & -3\sqrt{5}-5 \end{pmatrix}.$$
Entonces, si $q = \frac{p-2}{2}$, tenemos que
$$BVD^{\frac{p-2}{2}}V^{-1}Z_0^*A =$$
$$\frac{1}{5} \begin{pmatrix} 3+\sqrt{5} & 3-\sqrt{5} \\ 3+\sqrt{5} & 3-\sqrt{5} \end{pmatrix} \begin{pmatrix} \lambda_1^q & 0 \\ 0 & \lambda_2^q \end{pmatrix} \begin{pmatrix} 3\sqrt{5}-5 & 3\sqrt{5}-5 \\ -3\sqrt{5}-5 & -3\sqrt{5}-5 \end{pmatrix}$$
$$= \frac{4\sqrt{5}}{5} \begin{pmatrix} \lambda_1^q-\lambda_2^q & \lambda_1^q-\lambda_2^q \\ \lambda_1^q-\lambda_2^q & \lambda_1^q-\lambda_2^q \end{pmatrix}.$$

Entonces
$$\frac{4\sqrt{5}}{5} \begin{pmatrix} \lambda_1^q-\lambda_2^q & \lambda_1^q-\lambda_2^q \\ \lambda_1^q-\lambda_2^q & \lambda_1^q-\lambda_2^q \end{pmatrix} = 0,$$
si y sólo si $\lambda_1^q = \lambda_2^q$, si y sólo si $q = 0$ ó $p = 2$.

6.4 APLICACIONES DEL OPERADOR DE SHORTED

A continuación veremos aplicaciones y conclusiones que se desprenden de las secciones anteriores y prestaremos especial atención al operador $B^*W_{/R(A)}B$, el cual aparece naturalmente cuando estudiamos el Problema (6.4). Más precisamente, vimos que si $R(B) \subseteq R(A) + R(A)^{\perp_W}$, existe $X_0 \in L(\mathcal{H})$, tal que si $A \in CR(\mathcal{H})$, $B \in L(\mathcal{H})$, $W \in L(\mathcal{H})^+$ y $G : L(\mathcal{H}) \to L(\mathcal{H})$,

$$G(X) = (AX - B)^*W(AX - B),$$

entonces

$$\underset{X \in L(\mathcal{H})}{min} G(X) = G(X_0) = B^*W_{/R(A)}B. \tag{6.27}$$

Estudiaremos cuando el operador $B^*W_{/R(A)}B$ es efectivamente un operador de shorted y analizaremos resultados que se desprenden de este hecho.

Proposición 6.4.1. *Sea $A \in CR(\mathcal{H})$, $B \in L(\mathcal{H})$ y $W \in L(\mathcal{H})^+$, entonces*

$$inf\{G(X) : \ X \in L(\mathcal{H})\} = inf\{(Y - B)^*W(Y - B) : Y \in L(\mathcal{H}),\ R(Y) \subseteq R(A)\} =$$

$$= inf\{(EB)^*W(EB) : E^2 = E,\ N(E) = R(A)\} = B^*W_{/R(A)}B.$$

Demostración. Sea $R = \{Z = B - AX : X \in L(\mathcal{H})\}$, $S = \{Z = B - Y : Y \in L(\mathcal{H}),\ R(Y) \subseteq R(A)\}$, $T = \{Z = EB : E^2 = E,\ N(E) = R(A)\}$. Entonces $T \subseteq S \subseteq R$. De hecho, si $Z \in T$, entonces $Z = EB$, y $N(E) = R(I - E) = R(A)$.

Sea $V = (I - E)B$ entonces $R(V) = R((I - E)B) \subseteq R(A)$, y $Z = EB = -(V - B) \in S$, por lo tanto $T \subseteq S$. Por otro lado, sea $Z \in S$, entonces $Z = B - Y,\ R(Y) \subseteq R(A)$, entonces $Y = AA^\dagger Y$, por lo tanto $Z = B - A(A^\dagger Y) \in R$, entonces $S \subseteq R$. Entonces tenemos

$$B^*W_{/R(A)}B = inf\{G(X) : \ X \in L(\mathcal{H})\} = inf_{Z \in R} Z^*WZ \leq inf_{Z \in S} Z^*WZ =$$

$$= inf\{(Y - B)^*W(Y - B) : Y \in L(\mathcal{H}),\ R(Y) \subseteq R(A)\} \leq inf_{Z \in T} Z^*WZ =$$

$$= inf\{(EB)^*W(EB) : \ E^2 = E,\ N(E) = R(A)\} = B^*W_{/R(A)}B,$$

donde usamos el Lema 6.1.7 y la Proposición 6.1.8. □

Las siguientes observaciones resultan de aplicar el Lema 6.1.7, la Proposición 6.1.8 y la proposición anterior.

Observación 6.4.2. Sea $A \in CR(\mathcal{H}), B \in L(\mathcal{H}), W \in L(\mathcal{H})^+$ y $\mathcal{M} = B^{-1}(R(A))$. Entonces

$$Q = \{Z = BP_{\mathcal{M}}X - B : X \in L(\mathcal{H})\} \subseteq S \subseteq R.$$

De hecho, tomando $Y = BP_{\mathcal{M}}X$, tenemos que $R(Y) \subseteq R(A)$ y $Q \subseteq S$.

Observación 6.4.3. Sea $A \in CR(\mathcal{H}), B \in L(\mathcal{H}), W \in L(\mathcal{H})^+$ y $\mathcal{M} = B^{-1}(R(A))$. Entonces

$$B^*WB_{/\mathcal{M}} = inf\{(P_{\mathcal{M}}X - I)^*B^*WB(P_{\mathcal{M}}X - I) : \ X \in L(\mathcal{H})\} =$$
$$= inf\{(BP_{\mathcal{M}}X - B)^*W(BP_{\mathcal{M}}X - B) : \ X \in L(\mathcal{H})\} =$$
$$= inf_{Z \in Q} Z^*WZ \geq inf_{Z \in S} Z^*WZ = B^*W_{/R(A)}B.$$

El siguiente teorema, permite establecer que si B es de rango cerrado y $\mathcal{S} \subseteq \mathcal{H}$ es un subespacio cerrado, entonces el operador $B^*W_{/R(B)\cap\mathcal{S}}B$ es efectivamente un operador de shorted.

Teorema 6.4.4. *Sea $B \in CR(\mathcal{H})$, $W \in L(\mathcal{H})^+$ y $\mathcal{S}$ un subespacio cerrado, entonces*

$$B^*WB_{/B^{-1}(\mathcal{S})} = B^*W_{/R(B)\cap\mathcal{S}}B.$$

Demostración.

$$B^*WB_{/B^{-1}(\mathcal{S})} = inf\{(P_{B^{-1}(\mathcal{S})}X - I)^*B^*WB((P_{B^{-1}(\mathcal{S})}X - I) : \ X \in L(\mathcal{H})\} =$$
$$= inf\{(BP_{B^{-1}(\mathcal{S})}X - B)^*W(BP_{B^{-1}(\mathcal{S})}X - B) : \ X \in L(\mathcal{H})\} = B^*W_{/R(B)\cap\mathcal{S}}B,$$

donde usamos la Proposición 6.1.8 y el hecho que $R(BP_{B^{-1}(\mathcal{S})}) = BR(P_{B^{-1}(\mathcal{S})}) = BB^{-1}(\mathcal{S}) = R(B) \cap \mathcal{S}$, que es un subespacio cerrado. □

Corolario 6.4.5. *Sea $A, B \in CR(\mathcal{H})$, $W \in L(\mathcal{H})^+$, entonces*

$$B^*WB_{/B^{-1}(R(A))} = B^*W_{/R(B)\cap R(A)}B.$$

Corolario 6.4.6. *Sea $A \in CR(\mathcal{H})$, $B \in L(\mathcal{H})$ tal que $R(A) \subseteq R(B)$ y $W \in L(\mathcal{H})^+$, entonces*

$$B^*WB_{/B^{-1}(R(A))} = B^*W_{/R(A)}B.$$

Demostración. Siguiendo la prueba del Teorema 6.4.4, obtenemos la conclusión. □

Propiedades del operador $B^*W_{/R(A)}B$

A continuación estudiaremos algunas propiedades del operador $B^*W_{/R(A)}B$, el cual por las observaciones anteriores, es efectivamente un operador de shorted si por ejemplo $A \in CR(\mathcal{H})$ y $B \in L(\mathcal{H})$ tal que $R(A) \subseteq R(B)$ y $W \in L(\mathcal{H})^+$. Cuando $R(B) \subseteq R(A) + R(A)^{\perp_W}$, las propiedades que veremos a continuación, son las que se esperarían del operador $B^*W_{/R(A)}B$, si fuese efectivamente un operador de shorted, ver Capítulo 5.

Proposición 6.4.7. *Sea $A \in CR(\mathcal{H})$, $B \in L(\mathcal{H})$ y $W \in L(\mathcal{H})^+$. Si $R(B) \subseteq R(A) + R(A)^{\perp_W}$, entonces existe un operador $Z \in L(\mathcal{H})$ tal que*

$$Z^*WZ = B^*WZ = B^*W_{/R(A)}B.$$

Demostración. $R(B) \subseteq R(A) + R(A)^{\perp_W}$ si y sólo si $R(A^*WB) \subseteq R(A^*WA)$.

Sea $(A^*WA)^\dagger : D((A^*WA)^\dagger) = R(A^*WA) + R(A^*WA)^\perp \subseteq \mathcal{H} \to \mathcal{H}$, la inversa de Moore Penrose de A^*WA.

Entonces $V = (A^*WA)^\dagger(A^*WB) \in L(\mathcal{H})$.

Sea $Z = B - AV \in L(\mathcal{H})$, entonces

$$Z^*WZ = (B^* - V^*A^*)WZ = B^*WZ - V^*A^*W(B - AV) =$$

$$= B^*WZ - V^*A^*WB + V^*A^*WAV =$$

$$= B^*WZ - V^*A^*WB + V^*A^*WA(A^*WA)^\dagger(A^*WB) =$$

$$= B^*WZ - V^*A^*WB + V^*P_{\overline{R(A^*WA)}}A^*WB =$$

$$= B^*WZ - V^*A^*WB + V^*A^*WB = B^*WZ.$$

Finalmente, como

$$A^*W(AV - B) = A^*WA(A^*WA)^\dagger(A^*WB) - A^*WB =$$

$$= P_{\overline{R(A^*WA)}}A^*WB - A^*WB = 0,$$

por la Proposición 6.1.10, tenemos que

$$G(V) = (B - AV)^*W(B - AV) = Z^*WZ = B^*W_{/R(A)}B.$$

□

Observación 6.4.8. Sea $A \in CR(\mathcal{H})$, $B \in L(\mathcal{H})$ y $W \in L(\mathcal{H})^+$. Si existe un operador $Z \in L(\mathcal{H})$ tal que $Z^*WZ = B^*WZ$, entonces

$$R(B) \subseteq R(Z) + R(Z)^{\perp_W}.$$

De hecho, como Z es tal que $Z^*W(B - Z) = 0$, entonces $R(B - Z) \subseteq N(Z^*W) = W^{-1}(N(Z^*)) = W^{-1}(R(Z)^\perp) = R(Z)^{\perp_W}$. Sea $x \in \mathcal{H}$, entonces

$$Bx = Zx + Bx - Zx \in R(Z) + R(Z)^{\perp_W}.$$

Por lo tanto $R(B) \subseteq R(Z) + R(Z)^{\perp_W}$.

Teorema 6.4.9. *Sea $A \in CR(\mathcal{H})$, $B \in L(\mathcal{H})$ y $W \in L(\mathcal{H})^+$. Si $R(B) \subseteq R(A) + R(A)^{\perp_W}$, entonces*

$$N(B^*W_{/R(A)}B) = B^{-1}(N(W) + R(A)),\ R(B^*W_{/R(A)}B) \subseteq B^*(R(W) \cap R(A)^\perp)$$

$$y\ \overline{R(B^*W_{/R(A)}B)} = \overline{B^*(R(W) \cap R(A)^\perp)}.$$

Demostración. Sea $\mathcal{N} = \overline{W^{1/2}(R(A))}$, entonces si $R(B) \subseteq R(A) + R(A)^{\perp_W}$, tenemos que

$$\begin{aligned} &W^{1/2}B\mathcal{H} \subseteq W^{1/2}(R(A)) + W^{1/2}(W^{-1}(R(A)^{\perp})) = \\ &W^{1/2}(R(A)) + W^{-1/2}(R(A)^{\perp}) \cap R(W^{1/2}) = \\ &W^{1/2}(R(A)) + \mathcal{N}^{\perp} \cap R(W^{1/2}). \end{aligned} \tag{6.28}$$

Por otro lado

$$(W^{1/2}B)^{-1}(\mathcal{N} \cap R(W^{1/2})) = (W^{1/2}B)^{-1}(W^{1/2}(R(A))),$$

de hecho, sea $x \in (W^{1/2}B)^{-1}(\mathcal{N} \cap R(W^{1/2}))$, entonces $W^{1/2}Bx \in \mathcal{N} \cap R(W^{1/2})$. Por la ecuación (6.28), tenemos que

$$W^{1/2}Bx = W^{1/2}Ah + z,$$

donde $h \in \mathcal{H}$ y $z \in \mathcal{N}^{\perp} \cap R(W^{1/2})$, entonces $z = W^{1/2}Bx - W^{1/2}Ah \in \mathcal{N} \cap R(W^{1/2})$ y $z = 0$, por lo tanto $W^{1/2}Bx \in W^{1/2}(R(A))$ y $(W^{1/2}B)^{-1}(\mathcal{N} \cap R(W^{1/2})) \subseteq (W^{1/2}B)^{-1}(W^{1/2}(R(A)))$, la otra inclusión es trivial.

Finalmente, por el Teorema 5.1.6, tenemos que $N(B^*W_{/R(A)}B) = N(B^*W^{1/2}(1 - P_{\mathcal{N}})W^{1/2}B) = N((1 - P_{\mathcal{N}})W^{1/2}B) = (W^{1/2}B)^{-1}(\mathcal{N})$. Por lo tanto

$$N(B^*W_{/R(A)}B) = (W^{1/2}B)^{-1}(\mathcal{N}) = (W^{1/2}B)^{-1}(\mathcal{N} \cap R(W^{1/2})) =$$

$$\begin{aligned} &= (W^{1/2}B)^{-1}(W^{1/2}(R(A))) = B^{-1}(W^{-1/2}W^{1/2}(R(A))) = \\ &= B^{-1}(R(A) + N(W^{1/2})) = B^{-1}(R(A) + N(W)). \end{aligned}$$

Por otra parte, por la Proposición 6.4.7, tenemos que $B^*W_{/R(A)}B = Z^*WZ = B^*WZ$, donde $Z = B - AV$, con $V = (A^*WA)^{\dagger}A^*WB \in L(\mathcal{H})$.

Entonces $R(B^*W_{/R(A)}B) = R(Z^*WZ) = R(B^*WZ) = B^*R(WZ)$, donde $R(WZ) = R(WB - WAV) \subseteq R(A)^{\perp} = N(A^*)$, (ver prueba de Proposición 6.4.7). Por lo tanto $R(WZ) = R(WB - WAV) \subseteq R(W) \cap R(A)^{\perp}$. Entonces

$$R(B^*W_{/R(A)}B) = B^*R(WZ) \subseteq B^*(R(W) \cap R(A)^{\perp}).$$

Finalmente, por el Teorema 5.1.6, tenemos que $R(W) \cap R(A)^{\perp} \subseteq R(W_{/R(A)})$, entonces

$$B^*(R(W) \cap R(A)^{\perp}) \subseteq R(B^*W_{/R(A)}).$$

Además, vale que $\overline{R(B^*W_{/R(A)}B)} = N(B^*W_{/R(A)}B)^{\perp} = N(W_{/R(A)}B)^{\perp} = \overline{R(B^*W_{/R(A)})}$, entonces

$$B^*(R(W) \cap R(A)^{\perp}) \subseteq \overline{B^*(R(W) \cap R(A)^{\perp})} \subseteq \overline{R(B^*W_{/R(A)})} =$$

$$\overline{R(B^*W_{/R(A)}B)}.$$

fPor lo tanto
$$\overline{R(B^*W_{/R(A)}B)} = \overline{B^*(R(W) \cap R(A)^\perp)}.$$

□

A continuación veremos que propiedades debe cumplir el operador $W \in L(\mathcal{H})^+$ y el subespacio cerrado $\mathcal{M} \subseteq \mathcal{H}$, para que dado $B \in L(\mathcal{H})$, el operador $B^*W_{/\mathcal{M}}B$ sea efectivamente un operador de shorted.

Teorema 6.4.10. *Sea $\mathcal{M} \subseteq \mathcal{H}$, un subespacio cerrado, $B \in L(\mathcal{H})$ y $W \in L(\mathcal{H})^+$. Entonces*
$$B^*WB_{/\mathcal{M}} = B^*(P_{N(B^*)^\perp}WP_{N(B^*)^\perp})_{/\overline{B\mathcal{M}}}B.$$

Demostración.
En primer lugar, como siempre vale que $W_{/\overline{B\mathcal{M}}} \leq W$, tenemos que $B^*W_{/\overline{B\mathcal{M}}}B \leq B^*WB$, y como $R(B^*W_{/\overline{B\mathcal{M}}}B) = B^*R(W_{/\overline{B\mathcal{M}}}B) \subseteq B^*(B\mathcal{M}^\perp) = B^*[(B^*)^{-1}(\mathcal{M}^\perp)] = R(B^*) \cap \mathcal{M}^\perp \subseteq \mathcal{M}^\perp$, entonces vale que
$$B^*W_{/\overline{B\mathcal{M}}}B \leq B^*WB_{/\mathcal{M}}. \tag{6.29}$$

Caso 1: Supongamos $B \in GL(\mathcal{H})$. Entonces, sea $X \in L(\mathcal{H})$ tal que $R(X) \subseteq \mathcal{M}^\perp$ y
$$0 \leq X \leq B^*WB,$$
entonces conjugando la ecuación anterior por B^{-1} tenemos que
$$(B^*)^{-1}XB^{-1} \leq W.$$
Y si llamamos $Y = (B^*)^{-1}XB^{-1}$, entonces
$$R(Y) = (B^*)^{-1}(R(XB^{-1})) \subseteq (B^*)^{-1}(\mathcal{M}^\perp) = B\mathcal{M}^\perp.$$
Por lo tanto $Y \leq W_{/\overline{B\mathcal{M}}}$. Entonces
$$X = B^*YB \leq B^*W_{/\overline{B\mathcal{M}}}B.$$
Por lo tanto
$$B^*WB_{/\mathcal{M}} = max\{0 \leq X \leq B^*WB : \; R(X) \subseteq \mathcal{M}^\perp\} \leq B^*W_{/\overline{B\mathcal{M}}}B, \tag{6.30}$$
y usando (6.29), tenemos que
$$B^*WB_{/\mathcal{M}} = B^*W_{/\overline{B\mathcal{M}}}B = B^*(P_{N(B^*)^\perp}WP_{N(B^*)^\perp})_{/\overline{B\mathcal{M}}}B,$$
donde la última igualdad es trivial ya que en este caso $N(B^*)^\perp = \mathcal{H}$.
Caso 2: Supongamos que B^* es inyectivo. Llamaremos $(B^*)^{-1}$ al operador lineal no acotado, cuyo dominio es $R(B^*)$ y que cumple que para todo $x \in \mathcal{H}$ $(B^*)^{-1}B^*x = x$. Sea $X \in L(\mathcal{H})$ tal que $R(X) \subseteq \mathcal{M}^\perp$ y
$$0 \leq X \leq B^*WB.$$

Entonces por el Teorema de Douglas 2.1.1, existe $Z \in L(\mathcal{H})$ con $\|Z\| \leq 1$ tal que

$$X^{1/2} = B^* W^{1/2} Z.$$

Sea

$$Y^* = (B^*)^{-1} X^{1/2} = (B^*)^{-1} B^* W^{1/2} Z = W^{1/2} Z.$$

Entonces $Y^*Y = W^{1/2} Z Z^* W^{1/2} \leq W$, y por otra parte $R(Y^*Y) \subseteq R(Y^*) = R((B^*)^{-1} X^{1/2}) = (B^*)^{-1}(R(X^{1/2})) \subseteq (B^*)^{-1}(\overline{R(X)}) \subseteq (B^*)^{-1}(\mathcal{M}^\perp) \subseteq B\mathcal{M}^\perp$.

Por lo tanto $Y^*Y \leq W_{/\overline{B\mathcal{M}}}$, entonces

$$X = B^* Y^* Y B \leq B^* W_{/\overline{B\mathcal{M}}} B.$$

Por lo tanto

$$B^*WB_{/\mathcal{M}} = max\{0 \leq X \leq B^*WB : \ R(X) \subseteq \mathcal{M}^\perp\} \leq B^* W_{/\overline{B\mathcal{M}}} B,$$

y usando (6.29), tenemos que

$$B^*WB_{/\mathcal{M}} = B^* W_{/\overline{B\mathcal{M}}} B = B^*(P_{N(B^*)^\perp} W P_{N(B^*)^\perp})_{/\overline{B\mathcal{M}}} B,$$

donde la última igualdad es trivial ya que en este caso $N(B^*)^\perp = \mathcal{H}$.

Caso 3: Supongamos que $R(W) \subseteq N(B^*)^\perp = \overline{R(B)}$.

Sea $X \in L(\mathcal{H})$ tal que $R(X) \subseteq \mathcal{M}^\perp$ y

$$0 \leq X \leq B^*WB.$$

Entonces por el Teorema 2.1.1, existe $Z \in L(\mathcal{H})$ con $\|Z\| \leq 1$ tal que

$$X^{1/2} = B^* W^{1/2} Z.$$

Sea

$$Y^* = (B^*)^\dagger X^{1/2} = (B^*)^\dagger B^* W^{1/2} Z = P_{N(B^*)^\perp} W^{1/2} Z = W^{1/2} Z,$$

donde se usó que $R(W^{1/2}) \subseteq \overline{R(W)} \subseteq N(B^*)^\perp$.

Entonces $Y^*Y = W^{1/2} Z Z^* W^{1/2} \leq W$, y por otra parte $R(Y^*Y) \subseteq R(Y^*) = R((B^*)^\dagger X^{1/2}) = (B^*)^\dagger(R(X^{1/2})) \subseteq (B^*)^\dagger(\overline{R(X)}) \subseteq (B^*)^\dagger(\mathcal{M}^\perp) \subseteq B\mathcal{M}^\perp$, donde usamos que

$$B\mathcal{M}^\perp = (B^*)^{-1}(\mathcal{M}^\perp) = \{x \in \mathcal{H} : B^*x \in \mathcal{M}^\perp\} =$$

$$= \{x = x_1 + x_2 \in \mathcal{H} : x_1 \in N(B^*)^\perp,\ x_2 \in N(B^*),\ B^*x \in \mathcal{M}^\perp\} =$$

$$= \{x = (B^*)^\dagger B^* x + x_2 \in \mathcal{H} : x_2 \in N(B^*),\ B^*x \in \mathcal{M}^\perp\} =$$

$$= (B^*)^\dagger(\mathcal{M}^\perp) + N(B^*) \supseteq (B^*)^\dagger(\mathcal{M}^\perp).$$

Por lo tanto $Y^*Y \leq W_{/\overline{B\mathcal{M}}}$, entonces

$$X = B^* Y^* Y B \leq B^* W_{/\overline{B\mathcal{M}}} B.$$

Por lo tanto

$$B^*WB_{/\mathcal{M}} = max\{0 \leq X \leq B^*WB : \ R(X) \subseteq \mathcal{M}^{\perp}\} \leq B^*W_{/\overline{B\mathcal{M}}}B,$$

y usando (6.29), tenemos que

$$B^*WB_{/\mathcal{M}} = B^*W_{/\overline{B\mathcal{M}}}B.$$

Finalmente, vale que

$$B^*WB_{/\mathcal{M}} = B^*(P_{N(B^*)^{\perp}}WP_{N(B^*)^{\perp}})_{/\overline{B\mathcal{M}}}B,$$

pues, tomando $W' = P_{N(B^*)^{\perp}}WP_{N(B^*)^{\perp}}$, vale que $W' \geq 0$, y que $N(B^*) \subseteq N(W)$, entonces $R(W) \subseteq N(B^*)^{\perp}$ y por el caso anterior tenemos
$B^*WB_{/\mathcal{M}} = B^*W'B_{/\mathcal{M}} = B^*W'_{\overline{B\mathcal{M}}}B = B^*(P_{N(B^*)^{\perp}}WP_{N(B^*)^{\perp}})_{/\overline{B\mathcal{M}}}B.$ □

7

APLICACIONES

En este capítulo, luego de realizar una introducción teórica de algunas propiedades y definiciones que nos serán útiles, presentaremos problemas que surgen naturalmente en el área de Procesamiento de Imágenes y Señales, en los que se pueden aplicar los resultados obtenidos en el Capítulo 6.

Hacia el final del capítulo, presentaremos los llamados problemas de Procrusto, los cuales surgen naturalmente, por ejemplo, en Química Cuántica ó en la Teoría de Marcos.

7.1 OPTIMIZACIÓN CON RANGO FIJO

Sea $A \in \mathbb{C}^{m\times n}$, una matriz con $ran(A) = r$, (donde $ran(A)$ es el rango de la matriz A, es decir la dimensión del subespacio $R(A)$) llamaremos *descomposición en valores singulares* (DVS), a cualquier factorización

$$A = U\Sigma V^*, \tag{7.1}$$

con U y V unitarias y $\Sigma = \begin{pmatrix} D & 0_{r\times n-r} \\ 0_{m-r\times r} & 0_{m-r\times n-r} \end{pmatrix}$, con $D \in \mathbb{C}^{r\times r}$, una matriz diagonal con entradas diagonales positivas. Los *valores singulares* de A son las raíces cuadradas de los autovalores de A^*A y lo denotamos mediante $\sigma_i = \sqrt{\lambda_i}$. Vamos a suponer que los valores singulares de A están ordenados de forma decreciente.

Si llamamos u_i y v_i a las columnas i-ésimas de U y V respectivamente, se puede probar que $\{u_1, \cdots, u_r\}$, es una base ortonormal de $R(A)$, y $\{u_{r+1}, \cdots, u_m\}$, es una base ortonormal de $R(A)^\perp = N(A^*)$. Además $\|Av_i\| = \sigma_i$, para $1 \le i \le n$, y $\sigma_i = 0$ si, y sólo si, $i > r$, entonces $\{v_{r+1}, \cdots, v_n\}$ es una base ortonormal de $N(A)$ y $\{v_1, \cdots, v_r\}$ es una base ortonormal de $N(A)^\perp = R(A^*)$.

Cuando Σ contiene filas o columnas de ceros, es posible obtener una descomposición de A más compacta. Sea $r = ran(A)$, y dividamos U y V en las submatrices cuyos primeros bloques contengan r columnas: $U = [U_r\ U_{m-r}]$, con $U_r = [u_1 \cdots u_r]$ y $V = [V_r\ V_{n-r}]$, donde $V_r = [v_1\ \cdots\ v_r]$. Entonces, con la notación establecida

$$A = [U_r\ U_{m-r}]\ \Sigma\ [V_r\ V_{n-r}]^* = U_r D V_r^*. \tag{7.2}$$

Esta factorización de A se llama *descomposición en valores singulares reducida* de A.

La descomposición en valores singulares, es una herramienta que se utiliza, por ejemplo, para encontrar modelos de bajas dimensiones que aproximen en algún sentido, determinados datos, especialmente en el área de Señales y Procesamiento de imágenes [31, 42]. En particular, varios de estos problemas se pueden enunciar de la siguiente manera: dada una matriz $B \in \mathbb{C}^{n\times n}$ con $ran(B) \geq k$, para $k \in \mathbb{N}$ tal que $k < n$, hallar una matriz $Y_0 \in \mathbb{C}^{n\times n}$ con $ran(Y_0) = k$ tal que,

$$Y_0 = \underset{\{Y\in\mathbb{C}^{n\times n}:\ ran(Y)=k\}}{argmin} f(Y-B), \tag{7.3}$$

para alguna función de *costo* $f : \mathbb{C}^{n\times n} \to \mathbb{R}$.

Para los casos en que la función de costo f, viene dada por $|||\cdot|||$, es decir una norma unitariamente invariante, la solución a este problema es conocida. Supondremos que $ran(B) > k$, caso contrario, el mínimo se alcanza en $Y_0 = B$.

Estudiaremos tres casos del Problema (7.3). El primer caso, será para cuando f es la norma de Frobenius $\|\cdot\|_F$; el segundo caso para cuando f es la norma de operadores $\|\cdot\|$, y finalmente daremos la solución para el caso general de cuando f es una norma unitariamente invariante cualquiera (ver Sección 3.1.3).

Analicemos entonces, el siguiente problema:

$$Y_0 = \underset{\{Y\in\mathbb{C}^{n\times n}:\ ran(Y)=k\}}{argmin} \|Y-B\|_F, \tag{7.4}$$

Consideremos $B = U\Sigma V^*$, la descomposición en valores singulares de la matriz B.

Sea

$$B_k = \sum_{i=1}^{k} u_i \sigma_i v_i^*,$$

donde u_i es la i-ésima columna de la matriz unitaria U, v_i^* es la i-ésima fila de la matriz unitaria V^*, y σ_i el i-ésimo valor singular no nulo de B. Claramente, $ran(B_k) = k$.

Sea $Y \in \mathbb{C}^{n\times n}$, definimos la matriz (generalmente no diagonal) $D = U^*YV$, por lo tanto $Y = UDV^*$. Entonces

$$\|Y-B\|_F^2 = \|UDV^* - U\Sigma V^*\|_F^2 = \|D-\Sigma\|_F^2 =$$

$$\sum_i |\sigma_i - D_{ii}|^2 + \sum_{i\neq j} |D_{ij}|^2.$$

Entonces, la elección de D con $ran(D) = k$ que minimiza $\|Y-B\|_F^2$, es cuando D es diagonal, con los primeros k elementos de la diagonal (no nulos) iguales a $\sigma_1, \cdots, \sigma_k$ y los restantes $n-k$ elementos de la diagonal nulos. El Y correspondiente a dicho D es $Y_0 = B_k$, que resulta la solución de (7.4).

Ahora, supongamos que queremos hallar

$$Y_0 = \underset{\{Y \in \mathbb{C}^{n\times n}:\ ran(Y)=k\}}{argmin} \|Y - B\|, \tag{7.5}$$

donde $\|\cdot\|$ es la norma de operadores, es decir para $A \in \mathbb{C}^{m\times n}$,

$$\|A\| = \underset{x\in\mathbb{C}^n\ \|x\|=1}{sup} \|Ax\|.$$

En este caso, también resulta $Y_0 = B_k$.
Observar, en primer lugar que

$$\|B - B_k\| = \sigma_{k+1},$$

de hecho, sea $v \in \mathbb{C}^n$, tal que $\|v\| = 1$, entonces $v = \sum_{i=1}^{n} \alpha_i v_i$, con $\alpha_i \in \mathbb{C}$, $1 \leq i \leq n$, y v_i la i-ésima columna de la matriz unitaria V de la DVS de B. Entonces $\|v\|^2 = \sum_{i=1}^{n} |\alpha_i|^2 = 1$. Por lo tanto

$$\|(B - B_k)v\|^2 = \|\sum_{i=k+1}^{n}\sum_{j=1}^{n} \sigma_i u_i v_i^* \alpha_j v_j\|^2 = \|\sum_{i=k+1}^{n} \sigma_i \alpha_i u_i\|^2 =$$

$$= \sum_{i=k+1}^{n} \sigma_i^2 |\alpha_i|^2 \|u_i\|^2 = \sum_{i=k+1}^{n} \sigma_i^2 |\alpha_i|^2 \leq \sigma_{k+1}^2 \sum_{i=k+1}^{n} |\alpha_i|^2 \leq \sigma_{k+1}^2,$$

entonces

$$\|B - B_k\| \leq \sigma_{k+1}.$$

Y por otra parte, como

$$\|(B - B_k)v_{k+1}\| = \|\sum_{i=k+1}^{n} \sigma_i u_i v_i^* v_{k+1}\| = \|\sigma_{k+1} u_{k+1}\| = \sigma_{k+1},$$

tenemos que $\sigma_{k+1} \leq \|B - B_k\|$ y obtenemos la igualdad que queríamos.

Volviendo al Problema (7.5), sea $Y \in \mathbb{C}^{n\times n}$ tal que $ran(Y) = k$. Por el Teorema de la dimensión, tenemos que $dim(N(Y)) = n - dim(R(Y)) = n - k$. Sea $V_{k+1} = [v_1 \cdots v_{k+1}]$, es decir, la matriz que sólo contiene las primeras $k+1$ columnas de la matriz V de la DVS de B. Entonces

$$dim(R(V_{k+1})) + dim(N(Y)) = (k+1) + (n-k) = n+1,$$

entonces existe $x \in R(V_{k+1}) \cap N(Y)$, con $\|x\| = 1$. Sea $x = \sum_{i=1}^{k+1} c_i v_i$, con $c_i \in \mathbb{C}$, $1 \leq i \leq k+1$, tal que $\|x\|^2 = \|\sum_{i=1}^{k+1} c_i v_i\|^2 = \sum_{i=1}^{k+1} |c_i|^2 = 1$. Entonces, tenemos que,

$$\|Y - B\|^2 \geq \|(Y - B)x\|^2 = \|Bx\|^2 = \|\sum_{i=1}^{k+1} c_i B v_i\|^2 =$$

$$= \| \sum_{i=1}^{k+1} c_i \sigma_i u_i \|^2 = \sum_{i=1}^{k+1} |c_i|^2 \sigma_i^2 \|u_i\|^2 =$$

$$= \sum_{i=1}^{k+1} |c_i|^2 \sigma_i^2 \geq \sigma_{k+1}^2 \sum_{i=1}^{k+1} |c_i|^2 = \sigma_{k+1}^2 = \|B_k - B\|^2.$$

Por lo tanto $Y_0 = B_k$.

Finalmente, en [65], se estudia el problema más general de hallar

$$Y_0 = \underset{\{Y \in \mathbb{C}^{n\times n}:\ ran(Y) \leq k\}}{argmin} |||Y - B|||, \tag{7.6}$$

donde $||| \cdot |||$ es una norma unitariamente invariante (ver Sección 3.1.3). En este caso, también resulta $Y_0 = B_k$.

Para probar la existencia de solución del problema general (7.6), enunciaremos el siguiente lema, probado en [65, Lema 1].

Lema 7.1.1. *Sean $X, Y \in \mathbb{C}^{n\times n}$, con $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$ y $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_n$ los valores singulares de X e Y respectivamente. Si $\sigma_i \leq \alpha_i$, $1 \leq i \leq n$, entonces*

$$|||X||| \leq |||Y|||,$$

para cualquier $||| \cdot |||$, norma unitariamente invariante en $\mathbb{C}^{n\times n}$.

La siguiente notación será útil para resolver el problema general (7.6). Dada una matriz $X \in \mathbb{C}^{n\times n}$, para $1 \leq k \leq n$, denotaremos, $(X)_k$ a la matriz que se obtiene cuando cada elemento de las k últimas columnas de X se reemplaza por 0, y denotaremos por $[X]_k$ a la matriz que se obtiene cuando cada elemento de las k últimas filas y k columnas de X se reemplaza por 0. Si I es la matriz identidad, entonces es fácil ver que $(X)_k = X(I)_k$, $[X]_k = (I)_k X (I)_k$ y además $(XY)_k = X(Y)_k$. Denotaremos por $diag(x_1, \cdots, x_n)$ a la matriz de $\mathbb{C}^{n\times n}$ que tiene en su diagonal el vector $(x_1, \cdots, x_n)$ y 0 en el resto de sus entradas.

Usando el principio de min-max de Fischer-Courant [78, Pág. 21], es fácil probar que para cualquier matriz $X \in \mathbb{C}^{n\times n}$, tal que $X \geq 0$, vale que

$$\lambda_i([X]_k) \leq \lambda_i(X), \quad i = 1, \cdots, n, \tag{7.7}$$

$$\lambda_i([X]_k) \geq \lambda_{i-k}(X), \quad i = k+1, \cdots, n, \tag{7.8}$$

donde λ_i es el i-ésimo autovalor de X, tal que $\lambda_1 \leq \cdots \leq \lambda_n$.

Con está propiedad, dado $B \in \mathbb{C}^{n\times n}$, con $B = U\Sigma V^*$ su DVS, es posible probar que

$$|||(B)_k||| \leq |||B|||, \tag{7.9}$$

$$|||\Sigma - (\Sigma)_{n-k}||| = |||diag(0, \cdots, 0, \sigma_{k+1}, \cdots, \sigma_n)||| \leq |||(BP)_k|||, \qquad (7.10)$$

para toda norma unitariamente invariante $||| \cdot |||$ y para toda P unitaria.

De hecho, para probar (7.9), observar que $\{(B)_k\}^*(B)_k = [B^*B]_k$, entonces, por (7.7), los valores singulares de $(B)_k$ no son mayores a los valores singulares correspondientes de B. Por el Lema 7.1.1, tenemos (7.9).

Ahora, usando (7.8), tenemos para $k+1 \leq i \leq n$ y para cualquier P unitaria,

$$\lambda_i([P^*\Sigma^2 P]_k) \geq \lambda_{i-k}(P^*\Sigma^2 P) = \lambda_{i-k}(\Sigma^2) = \sigma^2_{n-(i-k)+1} = \lambda_i(\Sigma^2 - [\Sigma^2]_{n-k}).$$

Esta desigualdad, es válida para $1 \leq i \leq n$, pues $\lambda_i(\Sigma^2 - [\Sigma^2]_{n-k}) = 0$, cuando $1 \leq i \leq k$.

Más aún, tenemos que

$$\{\Sigma - (\Sigma)_{n-k}\}^*\{\Sigma - (\Sigma)_{n-k}\} = \Sigma^2 - [\Sigma^2]_{n-k},$$

y

$$\{(\Sigma P)_k\}^*(\Sigma P)_k = [P^*\Sigma^2 P]_k.$$

Por lo tanto, los valores singulares de $\Sigma - (\Sigma)_{n-k}$ no son mayores que los valores singulares correspondientes de $(\Sigma P)_k$, y por el Lema (7.1.1),

$$|||\Sigma - (\Sigma)_{n-k}||| \leq |||(\Sigma P)_k|||,$$

para cada P unitaria. Entonces, usando la DVS de B, tenemos

$$|||(BP)_k||| = |||B(P)_k||| = |||U\Sigma V^*(P)_k||| = |||(\Sigma V^* P)_k)||| \geq |||\Sigma - (\Sigma)_{n-k}|||,$$

para cada P unitaria y tenemos (7.10).

Volviendo al problema general (7.6), sea $Y \in \mathbb{C}^{n\times n}$, con $ran(Y) \leq k$. Entonces claramente, Y se puede expresar como $Y = P\Delta Q$ con $P,\ Q$ unitarias y $\Delta = diag(0, \cdots, 0, \delta_1, \cdots, \delta_k)$ (por ejemplo haciendo una permutación adecuada de la DVS de Y). Entonces, usando (7.9) tenemos

$$|||B - Y||| = |||(BQ^* - P\Delta)Q||| = |||BQ^* - P\Delta||| \geq |||(BQ^*)_k - (P\Delta)_k|||.$$

Pero $(P\Delta)_k = 0$, entonces usando (7.10) tenemos

$$|||B - Y||| \geq |||(BQ^*)_k||| \geq |||\Sigma - (\Sigma)_{n-k}|||,$$

para cada Y tal que $ran(Y) \leq k$.

Por otra parte, si $Y_0 = B_k = U(\Sigma)_{n-k}V^*$ entonces, tenemos que

$$B - Y_0 = B - B_k = U(\Sigma - (\Sigma)_{n-k})V^*$$

Por lo tanto:

$$|||B - Y_0||| = |||\Sigma - (\Sigma)_{n-k}|||.$$

Debido a su dificultad, el Problema (7.3), es estudiado usualmente relajando la restricción del rango de Y, lo cual bajo ciertas condiciones, resulta ser una relajación exacta. Para esto, se usa la factorización de $Y = AX$, con $A \in \mathbb{C}^{n\times k}, X \in \mathbb{C}^{k\times n}$.

Supongamos que la función *costo* viene dada por una norma unitariamente invariante $|||\cdot|||$, ahora estamos interesados en el siguiente problema:

$$Y_0 = \underset{\{X\in\mathbb{C}^{k\times n},\ A\in\mathbb{C}^{n\times k}\}}{argmin} |||AX - B|||. \tag{7.11}$$

De hecho, supongamos que $A_0 \in \mathbb{C}^{n\times k}, X_0 \in \mathbb{C}^{k\times n}$ satisface

$$|||A_0X_0 - B||| = \min_{\{X\in\mathbb{C}^{k\times n},\ A\in\mathbb{C}^{n\times k}\}} |||AX - B|||, \tag{7.12}$$

entonces,

$$|||A_0X_0 - B||| = \min_{X\in\mathbb{C}^{k\times n}} |||A_0X - B|||. \tag{7.13}$$

Luego $A_0X_0 = P_{R(A_0)}B$, (ver [40, Teorema 3.1] y [40, Proposición 2.5]). Entonces el problema (7.12) se puede reescribir como:

$$|||A_0X_0 - B||| = \min_{\{\mathcal{S}\subseteq\mathbb{C}^n:\ dim(\mathcal{S})\leq k\}} |||(I - P_{\mathcal{S}})B|||.$$

Observar que como

$$\{P_{\mathcal{S}}B - B : dim(\mathcal{S}) \leq k\} \subseteq \{Y - B : Y \in \mathbb{C}^{nxn},\ ran(Y) \leq k\},$$

entonces

$$\inf_{\{\mathcal{S}\subseteq\mathbb{C}^n:\ dim(\mathcal{S})\leq k\}} |||(I - P_{\mathcal{S}})B||| \geq \inf_{\{Y\in\mathbb{C}^{n\times n}:\ ran(Y)\leq k\}} |||B - Y||| = |||B - Y_0|||,$$

con $Y_0 = B_k$, donde se usó lo probado en (7.6).

Sean $U_1 \in \mathbb{C}^{n\times k}$ y $V_1 \in \mathbb{C}^{n\times k}$ las matrices que contienen las primeras k columnas de las matrices unitarias U y V de la DVS de B, respectivamente. Observar que $B_k = U_1\Sigma_kV_1^* = P_{R(U_1)}B$, donde Σ_k es la matriz que tiene en su diagonal los primeros k valores singulares de B: $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_k$. Por lo tanto

$$|||B - Y_0||| = |||B - P_{R(U_1)}B||| \geq \inf_{\{\mathcal{S}\subseteq\mathbb{C}^n:\ dim(\mathcal{S})\leq k\}} |||(I - P_{\mathcal{S}})B||| \geq |||B - Y_0|||.$$

De ésta manera, tenemos que el problema (7.11) tiene mínimo y el mismo se realiza en $Y_0 = P_{R(U_1)}B$.

Si un peso (positivo) se introduce en la ecuación (7.13), el Teorema 6.1.11, da condiciones suficientes y necesarias para la existencia de solución de dicho problema para $\|\cdot\|_p$.

7.2 TEORÍA DE MUESTREO Y RECONSTRUCCIÓN DE SEÑALES

En esta sección estudiaremos algunas aplicaciones de los problemas resueltos en la Sección 6.3, en primer lugar daremos algunas definiciones básicas y propiedades de la teoría de marcos.

7.2.1 *Marcos*

Sea $\mathcal{H}$ un espacio de Hilbert complejo y separable. Una sucesión $\mathcal{F} = \{f_j\}_{j\geq 1}$ es un marco ó *frame* para $\mathcal{H}$ si existen constantes $a, b > 0$ tales que

$$a\|f\|^2 \leq \sum_{j\geq 1} |\langle f, f_j \rangle|^2 \leq b\|f\|^2, \text{ para cada } f \in \mathcal{H}. \tag{7.14}$$

La mayor a y menor b con las propiedades de arriba se denominan las constantes del marco $\mathcal{F}$ óptimas y se denotan $a_{\mathcal{F}}$ y $b_{\mathcal{F}}$ respectivamente. Si $a_{\mathcal{F}} = b_{\mathcal{F}} = 1$, entonces $\mathcal{F}$ se llama un marco de *Parseval*. En líneas generales, el marco $\mathcal{F}$ es un conjunto generador de $\mathcal{H}$ donde permitimos combinaciones lineales de los elementos de $\mathcal{F}$ con coeficientes en $l^2(\mathbb{N})$. Más precisamente, dado un marco $\mathcal{F}$ definimos el operador de *sintésis* de $\mathcal{F}$ por $F : l^2(\mathbb{N}) \to \mathcal{H}$

$$F((\alpha_j)_{j\geq 1}) = \sum_{j\geq 1} \alpha_j f_j,$$

y el operador de *análisis* por $F^* : \mathcal{H} \to l^2(\mathbb{N})$,

$$F^* f = (\langle f, f_j \rangle)_{j\geq 1}.$$

Notar que la desigualdad de la derecha de (7.14) implica que F (y por lo tanto F^*) es un operador bien definido, lineal y acotado. Más aún, si llamamos $S_{\mathcal{F}} = FF^* \in L(\mathcal{H})$ entonces

$$S_{\mathcal{F}} f = FF^* f = \sum_{j\geq 1} \langle f, f_j \rangle f_j, \text{ para cada } f \in \mathcal{H}$$

y las desigualdades de (7.14) son equivalentes a las desigualdad de operadores

$$a \cdot I \leq S_{\mathcal{F}} \leq b \cdot I.$$

Por lo tanto, $S_F = FF^*$, el llamado operador del marco de $\mathcal{F}$, es un operador positivo e inversible y además vale que

$$a_{\mathcal{F}}^{1/2} = inf\{\|Fx\| : x \in N(F)^{\perp}, \|x\| = 1\} = \|F^{\dagger}\|^{-1}$$

y

$$b_{\mathcal{F}}^{1/2} = \|F\|.$$

Por otra parte, las identidades

$$f = S_{\mathcal{F}} S_{\mathcal{F}}^{-1} f = \sum_{j\geq 1} \left\langle S_{\mathcal{F}}^{-1} f, f_j \right\rangle f_j = \sum_{j\geq 1} \left\langle f, S_{\mathcal{F}}^{-1} f_j \right\rangle f_j, \text{ para cada } f \in \mathcal{H}, \tag{7.15}$$

muestran que $\mathcal{F}$ permite la representación lineal de vectores de $\mathcal{H}$, cuando la sucesión de coeficientes $\left\langle f, S_{\mathcal{F}}^{-1} f_j \right\rangle_{j\geq 1} = F^*(S_{\mathcal{F}}^{-1} f) \in l^2(\mathbb{N})$.

También notar que, si definimos $\mathcal{F}^\# = \{S_{\mathcal{F}}^{-1} f_j\}_{j\geq 1}$, entonces $\mathcal{F}^\#$ es un marco de $\mathcal{H}$ con operador del marco $S_{\mathcal{F}}^{-1}$. Esto muestra que los coeficientes dados por el mapa

$$\mathcal{H} \ni f \to (\left\langle f, S_{\mathcal{F}}^{-1} f_j \right\rangle)_{j\geq 1} \in l^2(\mathbb{N})$$

dependen de manera continúa en $f \in \mathcal{H}$, es decir, $\mathcal{F}$ permite fórmulas de reconstrucción lineales y estables.

Una importante propiedad de los marcos $\mathcal{F}$ es su redundancia. De hecho, es bien sabido que la falta de unicidad en este contexto es una ventaja, ya que provee herramientas para lidiar con problemas que aparecen en la aplicación de la teoría de marcos. Un medida típica de la redundancia del marco $\mathcal{F}$ está dada por el *exceso* de $\mathcal{F}$, que es la dimensión de $N(F)$. En caso que $\mathcal{F}$ tenga exceso no nulo, entonces existen infinitos marcos duales para $\mathcal{F}$, es decir, marcos $\mathcal{G} = \{g_j\}_{j\geq 1}$ tal que

$$f = \sum_{j\geq 1} \left\langle f, g_j \right\rangle f_j, \text{ para cada} f \in \mathcal{H}.$$

Notar que $\mathcal{F}^\#$ es un marco dual de $\mathcal{F}$, llamado el marco dual canónico.

Es fácil ver que que el conjunto de marcos duales para $\mathcal{F}$ está en biyección con el conjunto de operadores lineales y acotados $G : l^2(\mathbb{N}) \to \mathcal{H}$ tal que $FG^* = I_{\mathcal{H}}$, a través del mapa

$$G \to \mathcal{G} = \{Gf_j\}.$$

Este último hecho muestra que si $\mathcal{G}$ es un marco dual para $\mathcal{F}$, entonces $\mathcal{F}$ es un marco dual para $\mathcal{G}$ es decir

$$f = \sum_{j\geq 1} \left\langle f, g_j \right\rangle f_j = \sum_{j\geq 1} \left\langle f, f_j \right\rangle g_j, \text{ para cada} f \in \mathcal{H}.$$

En este caso, decimos que $(\mathcal{F}, \mathcal{G})$ es un par de marcos duales para $\mathcal{H}$.

En la Sección 7.3.4, se presentará un problema de Procrusto relacionado con la reconstrucción de señales derivada de fórmulas de reconstrucción para el par $(\mathcal{F}, \mathcal{G})$. Más precisamente, se buscarán marcos de Parseval $\mathcal{X} = \{x_j\}_{j\geq 1}$, que cumplen $XX^* = S_{\mathcal{X}} = I_{\mathcal{H}}$, que minimizan el peor caso de error de reconstrucción.

7.2.2 *Muestreo y Reconstrucción*

Sean $\{w_n\}_{n\geq 1} \subseteq \mathcal{H}$, el marco de un subespacio cerrado $\mathcal{R}$, denominado subespacio de reconstrucción y $\{v_n\}_{n\geq 1} \subseteq \mathcal{H}$ el marco de un subespacio cerrado $\mathcal{S}$, denominado subespacio de muestreo. Supongamos que, A y B son los operadores de sintesis correspondientes a los marcos $\{w_n\}_{n\geq 1}$ y $\{v_n\}_{n\geq 1}$ respectivamente, es decir $A, B : \ell^2(\mathbb{N}) \to \mathcal{H}$ son los operadores tales que, si $x = \{x_n\}_{n\in\mathbb{N}} \in \ell^2(\mathbb{N})$, $Ax = \sum_{n\geq 1} x_n w_n$ y $Bx = \sum_{n\geq 1} x_n v_n$, que son acotados dado que $\{v_n\}_{n\in\mathbb{N}}$ y $\{w_n\}_{n\in\mathbb{N}}$ son marcos y cumplen (7.14). Observar, que en este caso, tenemos que B^* es el operador de análisis del marco $\{v_n\}_{n\geq 1}$ y $B^* : \mathcal{H} \to l^2(\mathbb{N})$,

$$B^* f = (\langle f, v_n \rangle)_{n\geq 1}, \text{ para cada } f \in \mathcal{H}.$$

El proceso de muestreo y reconstrucción de un vector $f \in \mathcal{H}$ se resume en la aplicación sucesiva de dos operadores:

$$\hat{f} = AB^* f = A(\{\langle f, v_n \rangle\}_{n\geq 1}) = \sum_{n\geq 1} \langle f, v_n \rangle\, w_n.$$

Todo este proceso puede representarse como $\hat{f} = Qf$, donde Q es el operador lineal y acotado $Q : \mathcal{H} \to \mathcal{H}$, tal que

$$Q = AB^*. \tag{7.16}$$

Si el operador de sintésis A es un inverso por derecha del operador de análisis B^*, esto es si $B^*A = I_{l^2(\mathbb{N})}$, se dice que los operadores son *consistentes*.

Si los operadores B^* y A son consistentes, se verifica la noción de *muestreo consistente* propuesta en [82, 83], según la cual la función reconstruída debe producir las mismas muestras que la original:

$$B^*\hat{f} = B^*(AB^* f) = (B^*A)B^* f = B^* f.$$

Si los operadores B^* y A son consistentes, entonces el operador Q definido en (7.16) resulta idempotente. De hecho,

$$Q^2 = (AB^*)(AB^*) = A(B^*A)B^* = AB^* = Q.$$

Entonces Q resulta ser una proyección en $R(A) = \mathcal{R}$.

Si $\mathcal{S} \neq \mathcal{R}$ y los operadores B^* y A son consistentes, entonces Q define una proyección oblicua, de modo que $f - \hat{f} \in N(B^*) = R(B)^\perp = \mathcal{S}^\perp$, de hecho

$$B^*(f - \hat{f}) = B^* f - B^*AB^* f = 0.$$

Podemos interpretar que el error $f - \hat{f}$ es un elemento de $\mathcal{H}$ que el operador B^* no detecta.

Si $\mathcal{S} = \mathcal{R}$ y los operadores B^* y A son consistentes, resultará que $f - \hat{f} \in \mathcal{S}^\perp$ y Q define una proyección ortogonal, es decir $Q = P_{\mathcal{R}}$. En este caso, tendremos

$$\hat{f} = \underset{g\in\mathcal{S}}{arg\ min} \|f - g\|.$$

Para ampliar las posibilidades de lo hecho hasta ahora se propone un nuevo esquema general de muestreo y reconstrucción. En este esquema, las muestras obtenidas mediante la aplicación del operador de muestreo B^* pueden ser transformadas por un filtro digital (un operador $X \in L(l^2(\mathbb{N}))$) antes de la reconstrucción. Es decir,

$$\hat{f} = AXB^* f. \tag{7.17}$$

La función reconstruida no tiene por qué coincidir con la original. Sin embargo, resulta razonable exigir que si $f \in \mathcal{R}$ entonces $\hat{f} = f$. Esta exigencia implica aceptar la hipótesis que

$$\mathcal{R} \cap \mathcal{S}^\perp = \{0\}.$$

En efecto, si existiera una función no nula $g \in \mathcal{R} \cap \mathcal{S}^\perp$ resultaría $B^* g = 0$, pues $\mathcal{S}^\perp = R(B)^\perp = N(B^*)$, y no se lograría su reconstrucción exacta pese a ser un elemento de $\mathcal{R}$.

Además si agregamos la hipótesis de reconstrucción consistente y que $\mathcal{S} = \mathcal{R}$. Entonces la función original y reconstruída deberán tener las mismas muestras:

$$B^* f = B^* \hat{f} = B^* AXB^* f.$$

En este caso particular, X es tal que $Q = AXB^*$ resulta ser la proyección oblicua $Q = P_{\mathcal{R}//\mathcal{S}^\perp}$. De hecho, dada $f \in \mathcal{H}$, se cumple que $Q\hat{f} = AXB^* \hat{f} = AXB^* f = Qf$, con lo cual

$$Q^2 f = Q(Qf) = Q\hat{f} = Qf$$

y Q es una proyección.

Además $N(Q) \supseteq N(B^*) = \mathcal{S}^\perp$. Para probar la otra inclusión, sea $f \in N(Q)$ pero $f \notin \mathcal{S}^\perp$, entonces $B^* f \neq 0$, lo cual contradice que $B^* \hat{f} = B^* Q f = B^* 0 = 0$. Entonces $N(Q) = \mathcal{S}^\perp$.

Por otra parte, $R(Q) \subseteq R(A) = \mathcal{R}$. Dada $f \in \mathcal{S} = \mathcal{R}$, al ser $\hat{f}$ una reconstrucción consistente resulta $B^*(Qf - f) = 0$. Entonces $Qf - f \in N(B^*) = \mathcal{S}^\perp$ y como $Qf - f \in \mathcal{S}$, tenemos que $Qf = f$ y $f \in R(Q)$.

En el caso en que tenemos la hipótesis de reconstrucción consistente pero $\mathcal{S} \neq \mathcal{R}$, si se cumple que

$$\mathcal{R} \cap \mathcal{S}^\perp = \{0\},$$

entonces $Q = AXB^*$ resulta también ser la proyección oblicua $Q = P_{\mathcal{R}//\mathcal{S}^\perp}$. De hecho, sigue valiendo que $N(Q) = \mathcal{S}^\perp$ y que $R(Q) \subseteq \mathcal{R}$. Ahora, sea $f \in \mathcal{R}$, al ser $\hat{f}$ una reconstrucción consistente resulta $B^*(Qf - f) = 0$. Entonces $Qf - f \in \mathcal{S}^\perp \cap \mathcal{R}$, pero como $\mathcal{R} \cap \mathcal{S}^\perp = \{0\}$, entonces $Qf = f$ y $f \in R(Q)$.

En este caso en particular, la señal reconstruída $\hat{f}$ no necesariamente es una buena aproximación de f, dado que la distancia $\|f - \hat{f}\| = \|f - AXB^* f\|$ no está minimizada.

Ahora, independientemente del método utilizado de muestreo y reconstrucción, supongamos que queremos encontrar un filtro dígital $X \in \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$, tal que $AXB^* f$ sea una buena aproximación de f en $R(A) = \mathcal{R}$, es decir queremos que AXB^* aproxime a $P_{\mathcal{R}}$ en algún sentido. Por ejemplo, quisiéramos encontrar $X_0 : \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$ un operador lineal y acotado tal que, para cada $f \in \mathcal{H}$

$$\|(AX_0B^* - P_{\mathcal{R}})f\| \leq \|(AXB^* - P_{\mathcal{R}})f\|,$$

para cada $X \in L(\mathcal{H})$.

Dado $W \in L(\mathcal{H})^+$. Alternativamente y generalizando el problema anterior, quisieramos encontrar $X_0 : \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$ un operador lineal y acotado tal que, para cada $f \in \mathcal{H}$

$$\|(AX_0B^* - P_{\mathcal{R}})f\|_W \leq \|(AXB^* - P_{\mathcal{R}})f\|_W, \tag{7.18}$$

para cada $X \in L(\mathcal{H})$.

Esto significa que, estamos interesados en el siguiente problema,

$$\min_{X \in L(\ell^2(\mathbb{N}))} (AXB^* - P_{\mathcal{R}})^* W (AXB^* - P_{\mathcal{R}}), \tag{7.19}$$

con el orden inducido en $L(\mathcal{H})$ por el cono de operadores positivos.

En vista de lo probado en 6.3.3, cómo siempre vale que $R(P_{\mathcal{R}}) = \mathcal{R} \subseteq \mathcal{R} + \mathcal{R}^{\perp_W} = R(A) + R(A)^{\perp_W}$, el problema (7.19) tendrá solución si y sólo si

$$\mathcal{S}^{\perp} = N(B^*) \subseteq N(A^*WP_{\mathcal{R}}) = P_{\mathcal{R}}^{-1}W^{-1}(N(A^*)) =$$

$$= P_{\mathcal{R}}^{-1}W^{-1}(\mathcal{R}^{\perp}) = P_{\mathcal{R}}^{-1}(\mathcal{R}^{\perp_W}) = \mathcal{R}^{\perp} \oplus \mathcal{R} \cap \mathcal{R}^{\perp_W} = \mathcal{R}^{\perp} \oplus N(W) \cap \mathcal{R}.$$

Claramente, si tomamos $W = I$, (7.19) tendrá solución si y sólo si $\mathcal{S}^{\perp} \subseteq \mathcal{R}^{\perp}$.

El siguiente ejemplo, muestra como el peso W puede influir en la existencia de solución del problema (7.19).

Ejemplo 7.2.1. Sean $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$, con $W = I$. Entonces $\mathcal{S}^{\perp} = N(B^*) = gen\{(1,0)\}$ y como $\mathcal{R} = R(A) = gen\{(1,0)\}$, tenemos que $\mathcal{R}^{\perp} = R(A)^{\perp} = gen\{(0,1)\}$, por lo tanto $\mathcal{S}^{\perp} \not\subseteq \mathcal{R}^{\perp}$, y el problema (7.19) no tiene solución.

Si ahora tomamos, $W = B$. Entonces $N(W) = gen\{(1,0)\}$ y vale que $\mathcal{R}^{\perp} \oplus N(W) \cap \mathcal{R} = gen\{(0,1)\} \oplus gen\{(1,0)\} = \mathbb{R}^2$, y entonces $\mathcal{S}^{\perp} \subseteq \mathcal{R}^{\perp} \oplus N(W) \cap \mathcal{R}$, por lo tanto el problema (7.19) tiene solución.

Si $W \in L(\mathcal{H})^+$ también satisface que $W^{1/2} \in S_p$, la clase p-Schatten (para algún p con $1 \leq p < \infty$), también podríamos estar interesados en la existencia de

$$\underset{X \in L(\mathcal{H})}{min} \|AXB^* - P_{\mathcal{R}}\|_{p,W}. \tag{7.20}$$

En la segunda parte del capítulo 6, establecimos condiciones suficientes y necesarias para la existencia de solución del problema (7.20), tanto para el caso $p = 2$ como para el caso $1 \leq p < \infty$ y $N(B^*) \subseteq N(A^*WP_{\mathcal{R}}) = \mathcal{R}^{\perp} \oplus N(W) \cap \mathcal{R}$. Finalmente, para el caso general, fue posible dar condiciones suficientes para la existencia de dicho mínimo.

7.3 PROBLEMAS DE PROCRUSTO

En la mitología griega, Procrusto, también llamado Damastes, era un bandido y posadero del Ática (o según otras versiones a las afueras de Eleusis), se le consideraba hijo de Poseidón. Procrusto tenía su casa en las colinas, donde ofrecía posada al viajero solitario. Allí lo invitaba a tumbarse en una cama de hierro donde, mientras el viajero dormía, lo amordazaba y ataba a las cuatro esquinas del lecho. Si la víctima era alta, Procrusto la acostaba en una cama corta y procedía a serrar las partes de su cuerpo que sobresalían: los pies y las manos o la cabeza. Si por el contrario era baja, la invitaba a acostarse en una cama larga, donde también la maniataba y descoyuntaba a martillazos hasta estirarla (de aquí viene su nombre). Según otras versiones, nadie coincidía jamás con el tamaño de la cama porque ésta era secretamente regulable: Procrusto la alargaba o acortaba a voluntad antes de la llegada de sus víctimas.

En esta sección, decidimos presentar algunos ejemplos de problemas de Procrusto, los cuales se diferencian de los problemas que se resolvieron en este trabajo, en que tienen una condición adicional (restricción) que consta en general en buscar el mínimo pidiendo, por ejemplo, que el mismo resulte un operador unitario, etc. Si bien en la tesis no presentamos resultados para este tipo de problemas, nos pareció interesante incluir algunos ejemplos de este tipo de problemas de minimización.

En Química Cuántica ó en el área de Procesamiento de Señales, estos problemas de Procrusto surgen naturalmente. Realizaremos una breve introducción de algunos conceptos de Química Cuántica para entender la importancia de los problemas de Procrusto que se resolverán a continuación. Finalmente daremos un ejemplo de problema de Procrusto conmumente encontrado en la teoría de marcos y otro relacionado con el procesamiento y reconstrucción de señales.

7.3.1 *Problemas de Procrusto en Química Cuántica*

7.3.1.1 *Algunos conceptos introductorios de química cuántica*

Seguiremos la bibliografía [57] a fin de enunciar algunos conceptos básicos de la química cuántica. En Química Cuántica o Mecánica Cuántica, el estado de un sistema viene definido por una función matemática Ψ, llamada la función de estado o la función de onda dependiente del tiempo del sistema. Ψ es una función de las coordenadas de las partículas del sistema y también del tiempo. Por ejemplo, para un sistema de dos partículas tenemos, $\Psi = \Psi(x_1, y_1, z_1, x_2, y_2, z_2, t)$ donde (x_1, y_1, z_1) y $(x_2, y_2, z_2,)$ son las coordenadas de las partículas 1 y 2 respectivamente. Para un sistema con n partículas, la mecánica cuántica postula que la ecuación gobernando como varía Ψ con el tiempo t es:

$$-\frac{\hbar}{i}\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m_1}(\frac{\partial^2\Psi}{\partial x_1^2} + \frac{\partial^2\Psi}{\partial y_1^2} + \frac{\partial^2\Psi}{\partial z_1^2}) - \cdots - \frac{\hbar^2}{2m_n}(\frac{\partial^2\Psi}{\partial x_n^2} + \frac{\partial^2\Psi}{\partial y_n^2} + \frac{\partial^2\Psi}{\partial z_n^2}) + V\Psi, \tag{7.21}$$

donde $\hbar$ es la constante de Planck dividida por 2π; $m_1, \cdots, m_n$ son las masas de las partículas $1, \cdots, n$; (x_i, y_i, z_i) son las coordenadas espaciales de la partícula $1, \cdots, n$ y V es la energía potencial del sistema. Por supuesto V depende de las coordenadas de las partículas y podría variar con el tiempo.

La ecuación de Schrödinger independiente del tiempo

Para un átomo o molécula aislada, las fuerzas que actúan dependen sólo de las coordenadas de las partículas cargadas del sistema y son independientes del tiempo. Por lo tanto el potencial de energía V es independiente de t para un sistema aislado. Para los sistemas donde V es independiente del tiempo, la ecuación de Schrödinger independiente del tiempo tiene la forma $\Psi(x_1, \cdots, z_n, t) = f(t)\psi(x_1, \cdots, z_n)$, donde ψ es una función de las tres coordenadas de las partículas y f es una función del tiempo determinada. Esto es simple de demostrar tomando la ecuación de Schrödinger para una partícula en una dimensión. Si es independiente de t, tenemos:

$$-\frac{\hbar}{i}\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\Psi}{\partial x^2} + V\Psi. \tag{7.22}$$

Si proponemos una solución de la forma $\Psi(x,t) = f(t)\psi(x)$. Tenemos$\frac{\partial^2\Psi}{\partial x^2} = f(t)\frac{d^2\psi}{dx^2}$ y $\frac{\partial\Psi}{\partial t} = \psi(x)\frac{df}{dt}$. Sustituyendo en (7.22) y dividiendo por $f\psi = \Psi$ tenemos

$$-\frac{\hbar}{i}\frac{1}{f(t)}\frac{df}{\partial t} = -\frac{\hbar^2}{2m}\frac{1}{\psi(x)}\frac{\partial^2\psi}{\partial x^2} + V := E, \tag{7.23}$$

donde el parámetro E fue definido como

$$E := -\frac{\hbar}{i}\frac{1}{f(t)}\frac{df}{\partial t}. \tag{7.24}$$

La ecuación (7.23), muestra también que $-\frac{\hbar^2}{2m}\frac{1}{\psi(x)}\frac{\partial^2\psi}{\partial x^2} + V = E$. Por lo tanto E es independiente de t y de x y debe ser una constante. La constante E tiene las mismas dimensiones que V, por lo tanto tiene dimensiones de energía.

De hecho, la química cuántica postula a E como la energía del sistema. La ecuación (7.23) queda

$$-\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} + V(x)\psi(x) = E\psi(x), \tag{7.25}$$

que es la ecuación de Schrödinger independiente del tiempo para una partícula en una dimensión, además resolviendo (7.24), tenemos que

$$f(t) = e^{\frac{-iEt}{\hbar}}.$$

La ecuación (7.25) se puede resolver para ψ, conocido el potencial de energía $V(x)$. Entonces tenemos que

$$\Psi(x,t) = e^{\frac{-iEt}{\hbar}}\phi(x).$$

Para un sistema tridimensional y con n partículas, obtenemos la siguiente ecuación que tendremos que resolver para obtener la función ψ

$$-\frac{\hbar^2}{2m_1}\left(\frac{\partial^2\psi}{\partial x_1^2} + \frac{\partial^2\psi}{\partial y_1^2} + \frac{\partial^2\psi}{\partial z_1^2}\right) - \cdots - \frac{\hbar^2}{2m_n}\left(\frac{\partial^2\psi}{\partial x_n^2} + \frac{\partial^2\psi}{\partial y_n^2} + \frac{\partial^2\psi}{\partial z_n^2}\right) + V\psi = E\psi. \tag{7.26}$$

Con la ecuación (7.26), es posible determinar las ecuaciones de onda para una partícula en una caja ó una partícula libre en los casos unidimensional y tridimensional, la partícula en un pozo rectangular, el oscilador armónico etc., [56, Capítulo 2, 3 y 4].

La mecánica cuántica postula que no todas las funciones que satisfacen (7.26) son funciones de onda permitida. La solución que verifique (7.26) debe verificar que [56]:

1. ψ tiene un único y sólo un único valor en cada punto del espacio,
2. la función ψ debe ser continua y todas sus derivadas parciales también,
3. la integral en todo el espacio $\int |\psi|^2$, debe ser un número finito.

Las funciones de onda de cualquier operador que represente una magnitud física debe cumplir las condiciones anteriores. Cuando una función satisface dichas condiciones, se dice que se *comporta bien*.

Operador Hamiltoniano

Se define el operador Hamiltoniano, para un sistema de una partícula como:

$$\hat{H} = -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) + V(x,y,z),$$

definiendo el operador Laplaciano como $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$, obtenemos el operador Hamiltoniano tridimensional para una partícula.

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + V.$$

Para un sistema de n partículas tridimensional, obtenemos el siguiente operador:

$$\hat{H} = -\frac{\hbar^2}{2m_1}\nabla_1^2 - \cdots - -\frac{\hbar^2}{2m_n}\nabla_n^2 + V(x_1, \cdots, z_n).$$

De esta manera, la ecuación de Schrödinger independiente del tiempo se puede reescribir como:

$$\hat{H}\psi = E\psi. \tag{7.27}$$

Por lo tanto las ecuaciones de onda en (7.27) son autofunciones del operador Hamiltoniano $\hat{H}$, con los autovalores los valores de energía permitidos .

Dado que hay un conjunto de funciones de onda de estados estacionarios y de energías asociadas a estas, se puede escribir

$$\hat{H}\psi_j = E_j\psi_j,$$

donde j indica las distintas funciones de onda y sus energías asociadas.

Métodos de aproximación

Para sistemas de varios átomos o moléculas, los términos de repulsión interelectrónicas en la energía potencial V vuelven imposible resolver la ecuación de Schrödinger con exactitud. Se deben utilizar métodos de aproximación para dicha resolución. El método más usado, es el denominado método variacional. Sea $\hat{H}$ el operador Hamiltoniano independiente del tiempo de un sistema cuántico. De los postulados de la mecánica cuántica se puede deducir el siguiente teorema, cuya prueba se encuentra en [56, Capítulo 8]. Si ϕ es cualquier función que se *comporta bien* y normalizada de las coordenadas de las partículas del sistema, luego:

$$\int_{-\infty}^{\infty} \phi^* \hat{H}\phi d\tau \geq E_{gs}, \tag{7.28}$$

para ϕ normalizada. Donde E_{gs} es el estado fundamental de energía real del sistema y la integral se extiende en todo el espacio.

De hecho, tenemos que $\hat{H}\psi_n = E_n\psi_n$, para funciones de onda ψ_n y energías E_n, donde n denota los distintos estados de energía. Sea ϕ una función que se *comporta bien* y normalizada que depende de las coordenadas del sístema de partículas y sea $Z = \int_{-\infty}^{\infty} \phi^* \hat{H}\phi d\tau$.

Ya que el conjunto de autofunciones ψ_n es un conjunto completo [56, Postulado 4], podemos escribir $\phi = \sum_{j\geq 1} d_j\psi_j$, para ciertos $d_j \in \mathbb{C}$, donde $\phi^* = \sum_{j\geq 1} d_j^*\psi_j^*$. Luego como ϕ está normalizada, $\int_{-\infty}^{\infty} \phi^*\phi d\tau = 1$, tenemos que $\sum_{j\geq 1} |d_j| = 1$. Es fácil probar entonces que

$$Z = \sum_{j\geq 1} |d_i|^2 E_i \geq \sum_{j\geq 1} |d_i|^2 E_{gs} = E_{gs},$$

y podemos concluir que

$$Z = \int_{-\infty}^{\infty} \phi^* \hat{H}\phi d\tau \geq E_{gs}.$$

Para aplicar el método variacional, se toman diferentes funciones que se *comportan bien* y normalizadas $\phi_1, \phi_2, \cdots$, y para cada una de ellas se computa la integral variacional $\int_{-\infty}^{\infty} \phi^* \hat{H}\phi d\tau$. El teorema variacional, muestra que la función que nos da el valor más bajo de $\int_{-\infty}^{\infty} \phi^* \hat{H}\phi d\tau$ provee la aproximación mas cercana al estado fundamental de energía.

En caso de que ϕ no sea normalizada, tenemos que

$$\frac{\int_{-\infty}^{\infty} \phi^* \hat{H}\phi d\tau}{\int_{-\infty}^{\infty} \phi^*\phi d\tau} \geq E_{gs}. \tag{7.29}$$

La función ϕ se llama *función variacional de prueba*, y el cociente de integrales en (7.29), *integral variacional*. Luego ϕ, se puede usar como aproximación de la función de onda del estado fundamental real ψ_{gs} y se puede usar para aproximar el valor de E_{gs} real.

Una forma común de las funciones variacionales en mecánica cuántica son las funciones variacionales lineales

$$\phi = c_1 f_1 + \cdots + c_n f_n,$$

donde $f_1, \cdots, f_n$ son funciones y $c_1, \cdots, c_n$, son parámetros variacionales, cuyos valores se determinan mediante la minimización de la integral variacional. Sea

$$Y = \frac{\int_{-\infty}^{\infty} \phi^* \hat{H}\phi d\tau}{\int_{-\infty}^{\infty} \phi^*\phi d\tau},$$

luego las condiciones de mínimo para Y son

$$\frac{\partial Y}{\partial c_1} = \cdots = \frac{\partial Y}{\partial c_n} = 0.$$

Si nos limitamos a las funciones ϕ reales (de manera tal que los coeficientes c_i y las funciones f_i son todos reales) y definiendo $H_{jk} := \int_{-\infty}^{\infty} f_j^* \hat{H} f_k$, y $S_{jk} =:= \int_{-\infty}^{\infty} f_j^* f_k$, al igualar a 0 las derivadas parciales de Y, obtenemos las siguientes ecuaciones lineales

$$\sum_{k=1}^{n} [(H_{ik} - S_{ik}Y)c_k] = 0, \quad i = 1, 2, \cdots, n. \tag{7.30}$$

Este sistema de ecuaciones nos permiten hallar las constantes c_k. Observar que en (7.30), si queremos que la solución no sea la trivial (es decir $c_k = 0$, para $1 \leq k \leq n$), debemos pedir que el determinante de la matriz $[H_{ik} - S_{ik}Y]$ sea nulo.

El desarrollo del determinante de la matriz $H_{ik} - S_{ik}Y$, nos daría un polinomio de grado n en la variable Y, cuyas raíces, se puede demostrar, que son n y todas reales.

Si enumeramos dichas raíces de manera creciente, tenemos que

$$Y_1 \leq Y_2 \leq \cdots \leq Y_n,$$

si enumeramos los estados de energía en orden creciente, tenemos

$$E_{gs} \leq E_{gs+1} \leq \cdots \leq E_{gs+n} \leq \cdots,$$

donde E_{gs}, E_{gs+1}, son las energía del estado fundamental, el próximo estado menor, y así sucesivamente.

Resulta entonces que hay n diferentes conjuntos de $c_1, \cdots, c_n$ que satisfacen $\frac{\partial Y_i}{\partial c_1} = \cdots = \frac{\partial Y_i}{\partial c_n} = 0$, por lo que terminamos con n funciones variacionales distintas $\phi_1, \cdots, \phi_n$, y n valores diferentes para la ecuación variacional $Y_1, Y_2, \cdots, Y_n$, donde

$$Y_i = \frac{\int_{-\infty}^{\infty} \phi_i^* \hat{H} \phi_i d\tau}{\int_{-\infty}^{\infty} \phi_i^* \phi_i d\tau}.$$

Por lo que acabamos de ver, vale que $Y_1 \geq E_{gs}$, puede demostrarse también que $Y_2 \geq E_{gs+1}$, $Y_3 \geq E_{gs+2}$, etc.

De esta manera, se usan los valores de $Y_1, \cdots Y_n$, para aproximar las energías y las funciones de onda de los n estados de energía más bajo. Para aproximar a las energías de más estados, se agregarán más funciones f_k a la función de prueba ϕ.

Teoría de la perturbación

La teoría de aproximación por el método de perturbación consiste en lo siguiente. Sea $\hat{H}$ el operador Hamiltoniano independiente del tiempo, cuya ecuación de Schrödinger queremos resolver. En la teoría de aproximación por perturbación, se divide el operador $\hat{H}$ en dos partes:

$$\hat{H} = \hat{H}^0 + \hat{H}', \tag{7.31}$$

donde $\hat{H}^0$ es el operador Hamiltoniano cuya ecuación de Schrödinger puede ser resuelta exactamente y $\hat{H}'$ es un operador cuyo efecto se espera que sea bajo. El sistema con Hamiltoniano $\hat{H}^0$ se denomina sistema no perturbado, $\hat{H}'$ es llamada la perturbación y el sistema con Hamiltoniano $\hat{H} = \hat{H}^0 + \hat{H}'$ es denominado sistema perturbado.

En [56, Capítulo 9], se prueba que la energía E_n del estado n del sistema perturbado se puede escribir como:

$$E_n = E_n^{(0)} + E_n^{(1)} + E_n^{(2)} + \cdots ,$$

donde $E_n^{(0)}$ es la energía del estado del sistema no perturbado y $E_n^{(1)}, E_n^{(2)}, \cdots$, son llamadas las correcciones de primer orden, segundo orden, etc. a dicha energía.

Si el problema es apto para la teoría de las perturbaciones, las cantidades $E_n^{(1)}, E_n^{(2)}, \cdots$, decrecen con el orden considerado.

Para encontrar $E_n^{(1)}$ resolvemos la ecuación de Schrödinger $\hat{H}^0\psi_n^{(0)} = E_n^{(0)}\psi_n^{(0)}$ del sistema no perturbado. La teoría de la perturbación [56, Capítulo 9], nos muestra que el factor de corrección $E_n^{(1)}$ viene dado por:

$$E_n^{(1)} = \int_{-\infty}^{\infty} (\psi_n^{(0)})^* \hat{H}' \psi_n^{(0)} d\tau.$$

Dado que $\psi_n^{(0)}$ es conocido, $E_n^{(1)}$ se calcula fácilmente. En [56, Capítulo 9], se encuentran fórmulas equivalentes para $E_n^{(2)}, \cdots$, etc.

Como se puede ver en la breve introducción de algunos conceptos que se utilizan en la química cuántica, podemos observar la importancia de obtener conjuntos ortonormales de autofunciones para poder llevar a cabo aproximaciones que nos permitan resolver la ecuación de Schrödinger que no tienen solución analítica. Es muy importante también notar como el operador Hamiltoniano que es el que se utiliza para resolver la ecuación de Schrödinger es autoadjunto y además cuando se encuentran las autofunciones y autovalores asociados a dicho operador, las autofunciones de autovalores distintos son ortogonales y los autovalores todos son reales. Un aproximante al operador Hamiltoniano será conveniente que mantenga esta propiedad de ser autoadjunto. La posibilidad de reemplazar el operador Hamiltoniano infinito por uno finito que siga siendo autoadjunto, y por lo tanto encontrar una aproximación a E (y quizás ψ) con alguno de los procedimientos numéricos presentados en esta sección es uno de los objetivos del primer problema de aproximación que presentaremos. El segundo problema que presentaremos buscará ortogonalizar un conjunto de funciones lo que nos llevará a dilucidar el problema de Löwdin.

7.3.2 *Aproximación autoadjunta y el problema de Löwdin*

En esta sección resolveremos algunos de los problemas planteados en [61], a fin de aplicar los problemas de Procrusto a la Química Cuántica. El operador Hamiltoniano gobierna los sistemas de química cuántica. El Hamiltoniano, es un operador autoadjunto tal como vimos, en algún espacio de Hilbert de dimensión infinita. El ínfimo del espectro del hamiltoniano es a menudo un autovalor de multiplicidad 1. Este autovalor, y su correspondiente autovector (esencialmente único), son denominados los estados fundamentales de energía E_{gs} y la función de onda Ψ_{gs} fundamental del sistema. Encontrar (o preferentemente aproximar) E_{gs} y Ψ_{gs} es de gran importancia, dado que el estado fundamental de un sistema es el estado más estable y generalmente el estado más encontrado en la naturaleza. Dado que el Hamiltoniano es un operador de dimensión infinita, y dado que se tienen datos finitos solamente, se busca reemplazar el operador infinito original por uno finito, y por lo tanto encontrar una aproximación a E_{gs} (y quizás a Ψ_{gs}) con algún procedimiento numérico. Para ilustrar esto, supongamos que la molécula A_j consiste de N_j átomos ($j = 1, 2$) y estas 2 moléculas, A_1 y A_2 , se combinan para formar la molécula A_3 la cual tiene N_3 atomos. Uno desea obtener, una aproximación al estado de energía fundamental E_{A_3} y la función de onda Ψ_{A_3} del sistema A_3 a partir de los estados de energía y funciones de onda más simples E_{A_j}, Ψ_{A_j} del sistema constituido por A_j, $j = 1, 2$. Tal como vimos en la teoría de la perturbación en la sección de Química Cuántica. Por ejemplo la fórmula citada por Goldstein y Levy [49]

$$C_3H_8 = 2C_2H_6 - CH_4,$$

(la cual se lee como: un propano es igual a dos etanos menos un metano, sugiere la expresión del estado de energía fundamental $E_{C_3H_8}$ del propano en términos de $E_{C_2H_6}$ y E_{CH_4} del etano y metano, $E_{C_3H_8} = 2E_{C_2H_6} - E_{CH_4}$. Esta fórmula si bien no es exacta, representa bastante bien la realidad).

Supongamos ahora, que tenemos una matriz A de dimensión finita, la aproximación de algún Hamiltoniano pero que a diferencia del operador Hamiltoniano, no es autoadjunta (quizás por algún error numérico). Luego lo ideal sería reemplazarla por una matriz autoadjunta que es la *más cercana* a A en algún sentido. Veamos entonces, que el único aproximante (en S_2) autoadjunto de A es $Re(A)$. De hecho, si $X = X^*$ tenemos

$$\|A - Re(A)\|_2 = \|A - \frac{A + A^*}{2}\|_2 =$$

$$= \|\frac{1}{2}(A - X) + \frac{1}{2}(X - A^*)\|_2 \leq \frac{1}{2}\|A - X\|_2 + \frac{1}{2}\|X - A^*\|_2 = \|A - X\|_2,$$

donde se uso que $\|Z\|_2 = \|Z^*\|_2$.

Finalmente, la unicidad se sigue de la propiedad de convexidad de la norma $\| \cdot \|_2$, dado que el conjunto de matrices autoadjuntas es convexo, ver Lema 3.2.5.

En [59], podemos encontrar un resultado similar para S_p con $1 < p < \infty$, la prueba es idéntica que para el caso $p = 2$ y la unicidad del aproximante se sigue de la convexidad de las normas $\| \cdot \|_p$, probada en el Lema 3.2.5.

7.3.2.1 *El problema de Löwdin: aproximantes unitarios*

Cómo vimos en la sección anterior la estrategia para analizar un sistema en química cuántica es descomponer el sistema original en subsistemas más pequeños y luego recombinarlos en un sistema que de alguna forma se parece al original.

Supongamos que un sistema en Química Cuántica se descompuso en subsistemas y que esa información sobre los subsistemas es representada por una base no ortogonal $\{f_1, \cdots, f_n\}$ de vectores unitarios. En Química Cuántica todos los estados se representan por vectores unitarios. Con el objetivo de combinar este conjunto de vectores no ortogonales $\{f_1, \cdots, f_n\}$ para obtener información aproximada acerca del sistema original es necesario reemplazarlos por un conjunto de vectores unitarios $\{e_1, \cdots, e_n\}$, los cuales son ortogonales. Esto se lleva a cabo por medio de una ortogonalización, es decir buscaremos una matriz B, tal que:

$$Bf_i = e_i,$$

donde $\langle e_i, e_j \rangle = \delta_{ij}$, la delta de Kronecker. En otras palabras, B toma la base no ortogonal $\{f_1, \cdots, f_n\}$ y la transforma en una base ortogonal $\{e_1, \cdots, e_n\}$. Ahora, uno desea cambiar el conjunto $\{f_1, \cdots, f_n\}$ lo menos posible, porque cambiar las $\{f_1, \cdots, f_n\}$ resulta en una pérdida de información. En conclusión buscamos minimizar la distancia entre las dos bases $\{f_1, \cdots, f_n\}$ y $\{e_1, \cdots, e_n\}$, donde aquí interpretaremos distancia en el sentido de cuadrados mínimos, es decir, uno querría minimizar

$$\sum_i^n \|f_i - e_i\|^2.$$

Este procedimiento fue llevado a cabo por Löwdin en 1950 [58]. La ortogonalización de Löwdin es una ortogonalización que cumple que si la base dada $\{f_1, f_2, \cdots, f_N\}$, es permutada, para transformarse en $\{f_{\pi(1)}, f_{\pi(2)}, \cdots, f_{\pi(n)}\}$, entonces resulta el mismo B y la nueva base *más cercana* que obtenemos es $\{e_{\pi(1)}, e_{\pi(2)}, \cdots, e_{\pi(n)}\}$, donde $\{e_1, e_2, \cdots, e_N\}$, es la base ortonormal de Löwdin original (es decir $e_n = Bf_n, \ \ 1 \leq n \leq N$). Entonces, la ortonormalización de Löwdin difiere respecto del método de ortogonalización de Gram-Schmidt [44, Teorema 7.4], el cual si depende del orden en que se tomen los vectores para la ortogonalización.

En 1980 fue reformulado el problema de aproximación de Löwdin y resuelto como un problema de aproximación con matrices en vez de vectores [2]. Para obtener la reformulación, usaremos la descomposición polar de la matriz

$B = UL$, donde $L = |B|$ y U en este caso es una matriz unitaria y única pues B es inversible. Por lo tanto

$$\sum_{i=1}^{n} \|f_i - e_i\|^2 = \sum_{i}^{n} \|B^{-1}e_i - e_i\|^2 = \sum_{i=1}^{n} \|L^{-1}U^{-1}e_i - e_i\|^2 = \sum_{i=1}^{n} \|(L^{-1} - U)U^{-1}e_i\|^2 = \|L^{-1} - U\|_2^2. \tag{7.32}$$

La ultima igualdad en (7.32), se sigue del hecho que $\{U^{-1}e_i\}$ es una base ortonormal de $\mathcal{H}$ (ya que es unitario) y la definición de $\|\cdot\|_2$.

Para encontrar una matriz de ortogonalización que minimice la ecuación (7.32), reemplazamos B por otra matriz de ortogonalización, por ejemplo $\hat{B}$, y sea $\hat{B}f_i = \hat{e}_i$, $1 \leq i \leq n$, donde $\{\hat{e}_i\}_{1 \leq i \leq n}$ es otra base ortonormal de $\mathcal{H}$. Sea $Ve_i = \hat{e}_i$, $1 \leq i \leq n$. Luego V es una matriz unitaria de $\mathcal{H}$ a $\mathcal{H}$ y

$$f_i \underset{UL(=B)}{\rightarrow} e_i \underset{V}{\rightarrow} \hat{e}_i.$$

Luego, $\hat{B} = V(UL) = (VU)L$, dado que $\hat{B}$ es inversible, es la única descomposición polar de $\hat{B}$. Consecuentemente $L = |B| = |\hat{B}|$ es independiente de la matriz de ortogonalización, por lo que variar $B(= UL)$ sobre todas la matrices de ortogonalización B es equivalente a variar U sobre todas las matrices unitarias en $\mathcal{H}$ mientras mantenemos L fija.

Por la ecuación (7.32), debemos minimizar

$$\|L^{-1} - U\|_2$$

para una L^{-1} fija y positiva, variando U sobre todas las matrices unitarias en $\mathcal{H}$. Escribiendo $A = L^{-1}$, tenemos que: para una A fija y positiva queremos minimizar

$$\|A - U\|_2,$$

variando sobre el conjunto de las matrices unitarias en $\mathcal{H}$. Veamos entonces que el único aproximante unitario (en S_2) a una matriz positiva arbitraria y fija es la matriz identidad I. Sea $\{\phi_i\}$ una base ortonormal de $\mathcal{H}$. Entonces

$$\|A - U\|_2^2 = \sum_{i=1}^{n} \|(A - U)\phi_i\|^2 = \sum_{i=1}^{n} \|A\phi_i\|^2 + \|U\phi_i\|^2 - 2Re\left\langle A\phi_i, U\phi_i \right\rangle,$$

y se puede verificar que $\left\langle A\phi_i, U\phi_i \right\rangle$, está maximizado si $U = I$. Luego como

$$\|U\phi_i\| = \|\phi_i\| = 1,$$

tenemos

$$\|A - U\|_2^2 \geq \textstyle\sum_{i=1}^{n} \|A\phi_i\|^2 + \|\phi_i\|^2 - 2Re\left\langle A\phi_i, \phi_i \right\rangle = \|A - I\|_2^2. \tag{7.33}$$

Para la unicidad, no podemos usar la propiedad de convexidad de $\|\cdot\|_2$ ya que el conjunto de matrices unitarias no es convexo. Ahora, como I es un minimizante global, entonces es un extremo local de la función

$$F : U \to \|A - U\|_2,$$

con $U \in \mathcal{U}$, (llamamos $\mathcal{U}$ al conjunto de operadores unitarios). Indicamos que I debe ser el único extremo local, y por lo tanto el único minimizante global de F.

De hecho, si $V \in \mathcal{U}$ es un extremo local de F, luego la función $f : t \to F(e^{itP}V) = \|e^{itP}V - A\|_2$, donde $t \in \mathbb{R}$ y P es una proyección arbitraria de rango unidimensional, tiene un mínimo local en 0. Entonces, $f'(0) = 0 \leq f''(0)$. Si notamos $\sigma(V)$ al espectro de V, estas dos condiciones implican (ver [48]), dado que P es arbitrario, que $\sigma(V) \subseteq \mathbb{R}$ y $Re(\sigma(V)) \subseteq \mathbb{R}^+$. Entonces, como $V \in \mathcal{U}$, tenemos que $\sigma(V) = \{1\}$, lo que significa que $V = I$.

Extensión del resultado al espacio S_p

Dado $A \in L(\mathcal{H})^+$, sea $p \geq 1$, y

$$\mathcal{U}_A := \{U : U \text{ es unitaria}, U - A \in S_p\}.$$

En [2, Teorema 3.3], se prueba que si $U \in \mathcal{U}_A$, entonces $I \in \mathcal{U}_A$ y además ([2, Corolario 3.6])

$$\|U - A\|_p \geq \|I - A\|_p.$$

De esta manera, se extiende el resultado obtenido en (7.33) a los espacios S_p con $p \geq 1$ y de dimensión infinita. En este caso, se pierde la unicidad de I como minimizante. Con la hipótesis extra de que $A > 0$ (es estrictamente positiva), I es el único minimizante de $\|U - A\|_p$, $U \in \mathcal{U}_A$, para $p > 1$, [2, Teorema 3.5].

7.3.2.2 *Más problemas de aproximación tipo Löwdin*

Tal como hicimos con el problema de Löwdin, supongamos que tenemos una base arbitraria no ortogonal $\{f_1, \cdots, f_n\}$ de vectores unitarios. Supongamos que quisiéramos reemplazarlos con una base ortonormal $\{e_1, \cdots, e_n\}$ con las siguientes propiedades

1. $\{e_1, \cdots, e_n\}$ es cercano a $\{f_1, \cdots, f_n\}$,
2. $\{e_1, \cdots, e_n\}$ es cercano a otra base ortonormal $\{g_1, \cdots, g_n\}$.

La restricción 2., podría aparecer para probar una hipótesis seguida de alguna teoría química competitiva. Dado que 1. y 2. son restricciones que compiten entre sí, una solución óptima dependerá de cuanto peso relativo se les da a las restricciones 1. y 2. respectivamente. Para $0 \leq b \leq 1$, sea

$$F_b(B) = b\sum_i^n \|f_i - e_i\|^2 + (1-b)\sum_i^n \|g_i - e_i\|^2,$$

donde como en el problema de Löwdin, B es la matriz de ortogonalización tal que $Bf_i = e_i$, $1 \le i \le n$. Cómo antes, sea $B = UL$, donde $L = |B|$ y U es unitario, la única descomposición polar de la matriz B inversible . Tal como hicimos con el problema de Loẅdin, variar la matriz B es mantener su módulo L fijo mientras se varía su parte unitaria U.

Sea la matriz unitaria C que lleva la base $\{Lf_1, \cdots, Lf_n\}$ a la base ortonormal $\{g_1, \cdots, g_n\}$. Es decir $CLf_i = g_i$, $1 \le i \le n$. Luego $C^{-1}g_i = Lf_i$ y entonces $e_i = Bf_i = ULf_i = UC^{-1}g_i$. Entonces con $0 \le b \le 1$, tenemos

$$\begin{aligned} F_b(B) = b\sum_i^n \|(B^{-1} - I)e_i\|^2 + (1-b)\sum_i^n \|(I - UC^{-1})g_i\|^2 = \\ b\sum_i^n \|(L^{-1} - U)U^{-1}e_i\|^2 + (1-b)\sum_i^n \|(C - U)C^{-1}g_i\|^2 = \\ b\|L^{-1} - U\|_2^2 + (1-b)\|C - U\|_2^2. \end{aligned} \tag{7.34}$$

El problema ahora se transforma en el problema de minimizar la ecuación (7.34), variando la matriz unitaria U sujeta a la matriz positiva L^{-1}, con la matriz unitaria C y el número b ambos fijos.

En [48], se da una solución al problema presentado en (7.34), bajo ciertas hipótesis. Supongamos que $A_1, A_2, \cdots, A_N$ son operadores lineales en un espacio de Hilbert $\mathcal{H}$, N-dimensional. Sean $b_1 \ge b_2 \ge \cdots \ge b_n \ge 0$, con $b_1 + b_2 + \cdots + b_n = 1$. Supongamos que

$$D = \sum_{j=1}^n b_j A_j$$

es un operador normal. El problema de minimizar el funcional

$$\mathcal{E}(U) = \sum_{j=1}^n b_j \|U - A_j\|_2^2, \quad U \in \mathcal{U},$$

es equivalente a minimizar el funcional

$$\mathcal{F}(U) = \|U - D\|_2^2, \quad U \in \mathcal{U}.$$

Más aún, si escribimos

$$D = \sum_{j=1}^N \lambda_j e_j \otimes e_j, \tag{7.35}$$

donde $\{e_1, e_2, \cdots, e_N\}$ es una base ortonormal de $\mathcal{H}$ y $De_j = \lambda_j e_j$. Entonces, los operadores unitarios U que minimizan $\mathcal{E}$ tienen la siguiente forma. Para $1 \le j \le N$, $Ue_j = \mu_j e_j$, donde $\mu_j = \frac{\lambda_j}{|\lambda_j|}$, si $\lambda_j \ne 0$ y μ_j es un número complejo arbitrario de módulo 1 si $\lambda_j = 0$.

Por otra parte, si $\mathcal{N} = N(D)$. Entonces los operadores U que minimizan $\mathcal{E}$ son reducidos por $\mathcal{N}$ (es decir $U(\mathcal{N}) \subseteq \mathcal{N}$ y $U(\mathcal{N}^{\perp}) \subseteq \mathcal{N}^{\perp}$), son operadores unitarios arbitrarios en $\mathcal{N}$ y son operadores unitarios únicamente determinados en $\mathcal{N}^{\perp}$. Cuando D es inyectiva, el mínimo de $\mathcal{E}$ es alcanzado en un único U. Para las pruebas de estas afirmaciones, ver [48, Teorema 1].

Para el problema planteado en (7.34), tenemos que $D = bL^{-1} + (1-b)C$, entonces D será un operador normal si L^{-1} (ó equivalentemente L) conmutan con C. En ese caso, si D es de la forma (7.35), el U que minimiza (7.34) está dado por

$$Ue_j = \mu_j e_j,$$

para $j \in K = \{n \in \{1, 2, \cdots, N\} : \mu_j \neq 0\}$, donde $\mu_j = \frac{\lambda_j}{|\lambda_j|}$, y $\{Ue_j : j \in \{1, 2, \cdots, N\} \setminus K\}$ puede ser un base ortonormal arbitraria de $\{e_j : j \in K\}^{\perp}$.

7.3.3 *Aproximación simétrica de marcos y bases*

Dado un espacio de Hilbert $\mathcal{H}$ y un conjunto de vectores linealmente independiente $\{f_i\}_{i=1}^n$ en $\mathcal{H}$. El proceso de Gram-Schmidt es tradicionalmente usado para crear una base ortonormal de vectores $\{\mu_i\}_{i=1}^n$ a partir de $\{f_i\}_{i=1}^n$. De forma tal que

$$gen\{f_i\}_{1 \leq i \leq k} = gen\{\mu_i\}_{1 \leq i \leq k},$$

para cada $1 \leq k \leq n$. Este proceso depende del orden en el sentido en que, si se reordena el conjunto $\{f_i\}_{i=1}^n$, se obtiene un conjunto ortonormal totalmente nuevo $\{\mu_i\}_{i=1}^n$. En algunos problemas, es deseable tratar a los vectores $\{f_1, \cdots, f_n\}$ simultáneamente en un método que sea independiente del orden. Entonces, se busca ortonormalizar el conjunto $\{f_i\}_{i=1}^n$ de forma tal que

$$\sum_{i=1}^{n} \|u_i - f_i\|^2$$

es mínima, tal como vimos en la Sección 7.3.2. El conjunto resultante $\{\mu_i\}_{i=1}^n$ es llamado *simétrico* o la *ortogonalización de Löwdin* de $\{f_i\}_{i=1}^n$.

En [81], T. R. Tiballi, amplió los resultados vistos en la Sección 7.3.2 con un efoque de teoría de operadores. Investigó la existencia y unicidad de la ortogonalización simétrica de conjuntos $\{f_i\}_{i=1}^{\infty}$ linealmente independientes infinitos contables (finitos ó numerables) en espacios de Hilbert $\mathcal{H}$.

Considerando el operador $F : l^2(\mathbb{N}) \to \mathcal{H}$,

$$F(e_i) = f_i, \text{ para cada } i \in \mathbb{N}$$

y para la base ortonormal canónica $\{e_i\}_{i=1}^{\infty}$ de $l^2(\mathbb{N})$, probó la existencia de ortogonalizaciones simétricas de $\{f_i\}_{i=1}^{\infty}$, siempre que $(I - |F|)$ fuera un operador de Hilbert-Schmidt en $l^2(\mathbb{N})$ (ver B.1.2) y la dimensión del núcleo de F fuera menor o igual a la dimensión de $R(F)^{\perp}$. La unicidad se puede garantizar si y sólo si el núcleo de F es cero [81, Teorema 4].

En [43], se extienden estos resultados a marcos (ver Sección 7.2.1). En dicho trabajo, se estudia la existencia y unicidad de aproximaciones simétricas de marcos $\{f_i\}_{i=1}^{\infty}$ de subespacios de Hilbert $\mathcal{K} \subseteq \mathcal{H}$.

Diremos que dos marcos $\{f_i\}_{i\in\mathbb{N}}$ y $\{g_i\}_{i\in\mathbb{N}}$ de dos subespacios de Hilbert $\mathcal{K}$ y $\mathcal{L}$ de $\mathcal{H}$, respectivamente, son *débilmente similares*, si existe un operador lineal acotado e inversible $T : \mathcal{K} \to \mathcal{L}$ tal que $T(f_i) = g_i$ para cada índice i. Si $\mathcal{L} = \mathcal{K}$ entonces diremos que los marcos son *similares*.

Entonces, dado $\{f_i\}_{i\in\mathbb{N}}$, un marco de un subspacio $\mathcal{K}$ del espacio de Hilbert $\mathcal{H}$, un marco de Parseval $\{\nu_i\}_{i=1}^{\infty}$ en un subespacio de Hilbert $\mathcal{L} \subseteq \mathcal{H}$, se dice una *aproximación simétrica* de $\{f_i\}_{i\in\mathbb{N}}$, si los marcos $\{f_i\}_{i\in\mathbb{N}}$ y $\{\nu_i\}_{i\in\mathbb{N}}$ son débilmente similares, la suma

$$\sum_{j\geq 1} \|\nu_j - f_j\|^2$$

es finita y además es el mínimo de todas las sumas finitas $\sum_{j\geq 1} \|\mu_j - f_j\|^2$ que podrían aparecer para cualquier otro marco de Parseval $\{\mu_i\}_{i=1}^{\infty}$ de cualquier subespacio de Hilbert de $\mathcal{H}$ débilmente similar a $\{f_i\}_{i\in\mathbb{N}}$. Como resultado se obtiene que tal aproximación simétrica existe y es siempre única si y sólo si, el operador $(P - |F|)$ es de Hilbert-Schmidt, donde $(I - P) = P_{N(|F|)}$, [43, Teorema 2.3].

Dado $\{f_i\}_{i\in\mathbb{N}}$, un marco de un subspacio $\mathcal{K}$ del espacio de Hilbert $\mathcal{H}$, se define $F : l^2(\mathbb{N}) \to \mathcal{H}$ por

$$F(\sum_{j=1}^{\infty} \alpha_j e_j) = \sum_{j=1}^{\infty} \alpha_j f_j,$$

donde $\{e_i\}_{i\in\mathbb{N}}$ es la base canónica ortonormal de $l^2(\mathbb{N})$. Entonces F es el adjunto del operador de análisis y por lo tanto es acotado y tiene una descomposición polar $F = U|F|$, donde U es una isometría parcial de $l^2(\mathbb{N})$ en $\mathcal{H}$ con espacio inicial $N(F)^{\perp} \subseteq l^2(\mathbb{N})$ y rango $\overline{R(F)} \subseteq \mathcal{H}$ (ver A.1.6). Dado que el operador de $F^*|_{\mathcal{K}}$ es inyectivo y $F^* = U^*|F^*|$ es su descomposición polar, tenemos que $F|_{\mathcal{K}}$ tiene rango cerrado y que el conjunto $\{U(e_i)\}_{i\in\mathbb{N}}$ es un marco de Parseval de un subespacio separable de $\mathcal{H}$. El Teorema 2.3 de [43] muestra que, si $P = P_{R(F^*F)}$, denota la proyección de $l^2(\mathbb{N})$ en el subespacio cerrado $R(F^*F)$. Entonces el operador $(P - |F|)$ es de Hilbert-Schmidt si y sólo si la suma

$$\sum_{j\geq 1} \|\mu_j - f_j\|^2$$

es finita para al menos algún marco de Parseval $\{\mu_i\}_{i\in\mathbb{N}}$ de algún subespacio $\mathcal{L}$ de Hilbert de $\mathcal{H}$ que es débilmente similar a $\{f_i\}_{i\in\mathbb{N}}$. En este caso la desigualdad

$$\sum_{j=1}^{\infty} \|\mu_j - f_j\|^2 \geq \sum_{j=1}^{\infty} \|U(e_j) - f_j\|^2 = \|(P - |F|)\|_2,$$

vale para cualquier marco de Parseval $\{\mu_i\}_{i\in\mathbb{N}}$ de cualquier subespacio $\mathcal{L}$ de Hilbert de $\mathcal{H}$ que es débilmente similar a $\{f_i\}_{i\in\mathbb{N}}$. (La suma de la izquierda podría ser infinita para algunas elecciones de $\{\mu_i\}_{i\in\mathbb{N}}$). La igualdad aparece si y sólo si $\mu_i = U(e_i)$ para cada $i \in \mathbb{N}$. Consecuentemente, la aproximación simétrica de un marco $\{f_i\}_{i\in\mathbb{N}}$ en un espacio de Hilbert $\mathcal{K} \subseteq \mathcal{H}$ es un marco de Parseval que genera el mismo espacio de Hilbert $\mathcal{L} \equiv \mathcal{K}$ de $\mathcal{H}$ y que es similar a $\{f_i\}_{i\in\mathbb{N}}$.

7.3.4 *Problema de Procrusto en el procesamiento y reconstrucción de señales*

Siguiendo la notación y definiciones dadas en la Sección 7.2.1, daremos un ejemplo de problema de Procrusto hallado comunmente en el área de Procesamiento y Reconstrucción de Señales. Para ello, es necesario notar que, dado un marco $\mathcal{F}$ redundante, el marco dual canónico $\mathcal{F}^{\#}$ no es siempre la mejor elección para el dual de $\mathcal{F}$. Por ejemplo, la estabilidad numérica comienza a tener importancia cuando se lidia con fórmulas de reconstrucción derivadas del par dual $(\mathcal{F}, \mathcal{G})$; en este caso una medida de la estabilidad de las fórmulas de reconstrucción para un $\mathcal{F}$ fijo, vene dada por el *número de condición* κ del operador $S_{\mathcal{G}}$ de $\mathcal{G}$, definido por

$$\kappa(S_{\mathcal{G}}) = \|S_{\mathcal{G}}^{-1}\| \;\; \|S_{\mathcal{G}}\| \geq 1.$$

En [51], Han caracteriza los marcos $\mathcal{F}$ para los cuales el número de condición es mínimo, denotándolos por $\mathcal{X} = \{x_j\}_{j\geq 1}$, el cual es un marco de Parseval y en este caso $S_{\mathcal{X}} = I_{\mathcal{H}}$ y por lo tanto el número de condición de $S_{\mathcal{X}}$ es mínimo. Las condiciones en [51], para la existencia de un marco dual de Parseval de un marco $\mathcal{F} = \{f_j\}_{j\geq 1}$ resultan ser equivalentes a las condiciones para la existencia de un espacio de Hilbert mayor $\mathcal{K} \supset \mathcal{H}$, una base ortonormal $\{k_j\}_{j\geq 1}$ de $\mathcal{K}$ y una proyección oblicua $Q : \mathcal{K} \to \mathcal{H}$ tal que

$$Qk_j = f_j, \text{ para cada } j \in \mathbb{N},$$

las cuales fueron halladas en [6].

En el caso en que el marco $\mathcal{F}$ no admite un marco dual de Parseval, entonces se pueden considerar dos alternativas: se pueden buscar marcos duales que son óptimos para la estabilidad numérica ó se pueden buscar marcos de Parseval $\mathcal{X}$, que son óptimamente estables, que minimzan el error de reconstrucción cuando se derivan fórmulas de reconstrucción para el par $(\mathcal{F}, \mathcal{X})$. En [29], se considera la segunda alternativa y se buscan los marcos de Parseval $\mathcal{X}$ que minimizan el error del peor caso del algoritmo de reconstrucción derivado del par $(\mathcal{F}, \mathcal{X})$. Explícitamente, se buscan los marcos de Parseval $\mathcal{X} = \{x_j\}_{j\geq 1}$, es decir, aquellos que cumplen $XX^* = S_{\mathcal{X}} = I_{\mathcal{H}}$, que minimizan el peor caso de error de reconstrucción.

En [29], la medida del error de reconstrucción que se considera es $|||FX^* - I|||$, donde $||| \cdot |||$ es cualquier norma unitariamente invariante y se define

$$\alpha_{|||\cdot|||}(\mathcal{F}) = inf\{|||FX^* - I|||,\ XX^* = I_{\mathcal{H}}\},$$

con F y X los operadores de síntesis de $\mathcal{F}$ y $\mathcal{X}$ respectivamente. Se puede probar que $\alpha_{|||\cdot|||}(\mathcal{F})$, depende del espectro del operador $S_{\mathcal{F}}$ y también del exceso de $\mathcal{F}$. En el caso en que $\alpha_{|||\cdot|||}(\mathcal{F})$ se alcanza para algún marco de Parseval $\mathcal{X}$, se introduce el conjunto de marcos de Parseval cuasi-duales, definido, por

$$\Xi(\mathcal{F}) = \{\mathcal{X} = \{x_j\}_{j\geq 1} : XX^* = I,\ \alpha(F) = |||FX^* - I|||\}.$$

En estos casos, $\alpha_{|||\cdot|||}(\mathcal{F})$ es el peor error del algoritmo de reconstrucción basado en el par óptimo $(\mathcal{F}, \mathcal{X})$, donde $\mathcal{X}$ es un marco de Parseval cuasi-dual de $\mathcal{F}$.

8 CONCLUSIONES

En la Sección 6.3.2, dados $A,\ B \in CR(\mathcal{H})$, $C \in L(\mathcal{H})$ y $W \in L(\mathcal{H})^+$, estudiamos el problema de analizar la existencia de

$$\min_{X \in L(\mathcal{H})} \|AXB - C\|_{p,W}, \tag{8.1}$$

donde $\|X\|_{p,W} = \|W^{1/2}X\|_p$. Para obtener los resultados resumidos en el Teorema 6.3.9, supusimos como hipótesis que $W \in L(\mathcal{H})^+$ sea un operador tal que $W^{1/2} \in S_p$, para algún $1 \leq p < \infty$.

Observar que, dados $A,\ B \in CR(\mathcal{H})$, y $W \in L(\mathcal{H})^+$. Si definimos el conjunto

$$\Delta_p = \{X \in L(\mathcal{H}) : W^{1/2}(AXB - C) \in S_p\},$$

para algún p, con $1 \leq p < \infty$ y consideramos $F_p : \Delta_p \to \mathbb{R}$,

$$F_p(X) = \|W^{1/2}(AXB - C)\|_p,$$

en los casos en que el conjunto $\Delta_p \neq \emptyset$, obtendremos las mismas conclusiones que las obtenidas en el Teorema 6.3.9, sin pedir que $W^{1/2}$ pertenezca a alguna clase de Schatten p.

Más aún, si pensamos el problema estudiado en la Sección 6.1.2, como un caso particular del problema estudiado en la Sección 6.3.2; un mismo razonamiento se puede hacer, para obtener soluciones de dicho problema sin pedir que el operador $W \in L(\mathcal{H})^+$ sea un operador tal que $W^{1/2} \in S_p$, para algún $1 \leq p < \infty$.

En la Sección 6.3.1, se estudió el Problema de analizar la existencia de

$$\min_{X \in L(\mathcal{H})} (AXB - C)^* W (AXB - C), \tag{8.2}$$

con el orden inducido por el cono de operadores positivos en $L(\mathcal{H})$, que es una generalización del problema estudiado en 6.1.2, en ese caso $B = I$. En ambos casos se obtuvo que la existencia de este mínimo es equivalente a que se verifiquen dos condiciones: $R(C) \subseteq R(A) + R(A)^{\perp_W}$ y $N(B) \subseteq N(A^*WC)$ (para el caso en que $B = I$, la segunda condición se cumple trivialmente).

La Proposición 3.1.15, permite extender los resultados del mínimo obtenido para el orden inducido por el cono de operadores positivos en $L(\mathcal{H})$, a cualquier norma unitariamente invariante. Es decir, resolver el problema del mínimo (8.2), permite obtener condiciones suficientes para la existencia del mínimo del Problema (8.1) para cualquier norma unitariamente invariante, por ejemplo, las normas p con $1 \leq p < \infty$ ó la norma de operadores.

Para establecer condiciones necesarias para la existencia del mínimo de

$$\min_{X \in L(\mathcal{H})} \|AX - B\|_{p,W}, \tag{8.3}$$

para las normas p con $1 \leq p \leq \infty$, se utilizaron las fórmulas de las derivadas ϕ-direccionales de los operadores $\|X\|_p^p$, para $1 \leq p \leq \infty$ (ver Sección 3.2.1). En este caso, la idea fue probar que un mínimo global del Problema (8.3) es solución de la ecuación normal $A^*W(AX - B) = 0$.

Para el caso $1 \leq p < \infty$, fue posible probar que todo mínimo global de (8.3), es efectivamente una solución de la ecuación normal $A^*W(AX - B) = 0$, estableciendose una equivalencia entre la existencia de mínimo del problema asociado en el orden inducido por el cono de operadores positivos y el mínimo para las normas p con $1 \leq p < \infty$. Para el caso particular $p = \infty$, se pudo probar que un mínimo $X_0 \in L(\mathcal{H})$ del Problema (8.3) cumple que $A^*W(AX_0 - B)f = 0$, para algún f en el subespacio donde el operador $|W^{1/2}(AX_0 - B)|$ alcanza su norma.

Para el caso del mínimo del problema

$$\min_{X \in L(\mathcal{H})} \|AXB - C\|_{p,W}, \tag{8.4}$$

fue posible encontrar un equivalencia con el mínimo del problema asociado en el orden de operadores, cuando se cumple la condición $N(B) \subseteq N(A^*WC)$; es decir, el mínimo del Problema (8.4), es solución de la ecuación normal esperada $A^*W(AXB - C) = 0$ en este caso particular. Se puede ver fácilmente, que si se cumple la condición $N(B) \subseteq N(A^*WC)$, entonces la ecuación normal $A^*W(AXB - C) = 0$ es equivalente a $A^*W(AXB - C)B^* = 0$.

En particular, para el caso $p = 2$ y sin imponer restricciones al $N(B)$, se encontró una equivalencia entre el mínimo del Problema (8.4) en normas p y la existencia de solución de la ecuación normal $A^*W(AXB - C)B^* = 0$.

Para el caso general, $1 \leq p < \infty$ y sin imponer restricciones en el $N(B)$, fue posible encontrar ejemplos donde había mínimo del Problema (8.4) en las normas p pero no así en el orden de operadores (pues $N(B) \not\subseteq N(A^*WC)$), más aún, para $p \neq 2$, el mínimo no era solución de la ecuación normal $A^*W(AXB - C)B^* = 0$ (ver Ejemplo 2). También se encontraron ejemplos donde las soluciones de la ecuación normal $A^*W(AXB - C)B^* = 0$, no realizaban el mínimo del problema correspondiente en las normas p (ver Ejemplo 3), estos hechos contradicen el Teorema 4.1 de [63], por lo cual, los resultados obtenidos en este trabajo son nuevos incluso para matrices.

En la Sección 6.2, se usaron los resultados encontrados para resolver el Problema (8.3), para determinar, entre todos los operadores que minimizan dicho problema, aquellos de norma mínima.

Dado que los operadores que minimizan el Problema (8.3), son de la forma $X_0 = B' + P_{N(A^*WA)}Z$, donde $B' = (A^*WA)^\dagger A^*WB$ y $Z \in L(\mathcal{H})$, resolver

$$\min_{X \in M_B} \|X\|_{q,W_2}, \tag{8.5}$$

con M_B el conjunto de los X que minimizan (8.3), $W_2^{1/2} \in L(\mathcal{H})^+$, tal que $W_2 \in S_q$ para algún q, con $1 \leq q < \infty$, se vuelve equivalente a resolver el Problema (8.3) y todos los resultados obtenidos para ese caso se pueden usar.

En la Sección 6.4, prestamos especial antención al operador $B^*W_{/R(A)}B$, el cual aparece naturalmente cuando estudiamos el Problema (8.3). Más precisamente, vimos que si $R(B) \subseteq R(A) + R(A)^{\perp_W}$, entonces existe $X_0 \in L(\mathcal{H})$, tal que

$$(AX_0 - B)^*W(AX_0 - B) = B^*W_{/R(A)}B.$$

Comparando el Teorema 6.4.9 con la Proposición 5.2.1, se puede observar que cuando $R(B) \subseteq R(A) + R(A)^{\perp_W}$, hay cierta similitud entre $N(B^*W_{/R(A)}B)$ y $R(B^*W_{/R(A)}B)$ respecto al $N(W_{/R(A)})$ y $R(W_{/R(A)})$ cuando el par $(W, R(A))$ es compatible.

En el Corolario 6.4.6, se prueba que si $R(A) \subseteq R(B)$, entonces el operador $B^*W_{/R(A)}B$ resulta efectivamente un operador de shorted.

Por último, en el Teorema 6.4.10, dado $\mathcal{M} \subseteq \mathcal{H}$ un subespacio cerrado, $B \in L(\mathcal{H})$ y $W \in L(\mathcal{H})^+$, se pudo probar que cierto operador shorted conjugado por el operador B es efectivamente un operador shorted, más precisamente, se demostró que

$$B^*WB_{/\mathcal{M}} = B^*(P_{N(B^*)^\perp}WP_{N(B^*)^\perp})_{/\overline{B\mathcal{M}}}B.$$

Anexos

A

GENERALIDADES SOBRE ESPACIOS DE HILBERT

En este anexo se realiza un repaso de las propiedades más importantes de los espacios de Hilbert, se establece el Teorema espectral, el Teorema de cálculo funcional y el Teorema de diagonalización de operadores compactos normales.

A.1 ESPACIOS DE HILBERT

Sean X, Y dos espacios normados sobre $\mathbb{C}$; un operador lineal $T : X \to Y$ es continuo si y sólo si es acotado, es decir que el conjunto de números reales $\{\|Tx\| : x \in X, \|x\| \leq 1\}$ es acotado superiormente. El conjunto $L(X,Y)$ de los operadores líneales de X en Y es un espacio vectorial sobre $\mathbb{C}$, y tiene una norma definida por

$$\|T\| = sup\{\|Tx\| : x \in X, \|x\| \leq 1\}.$$

Si Y es un espacio de Banach, entonces $L(X,Y)$ también lo es.

Un espacio de Hilbert $\mathcal{H}$, es un espacio vectorial provisto de un producto interno al que notaremos $\langle \cdot, \cdot \rangle$, que con la distancia definida por dicho producto interno es completo.

Sea $\mathcal{H}$ un espacio de Hilbert, dados $x \in \mathcal{H}$ y $r > 0$, definimos la bola abierta con centro en x y radio r como

$$B(x,r) = \{y \in \mathcal{H} : \|y - x\| < r\}.$$

Dado un conjunto $Z \subseteq \mathcal{H}$, definimos el interior de Z como

$$Z^\circ = \{x \in Z : \exists r > 0 : B(x,r) \subseteq Z\},$$

y definimos la clausura Z como

$$\overline{Z} = \{x \in H : B(x,r) \cap Z \neq \emptyset, \text{ para cada } r \in \mathbb{R}\}.$$

La siguiente equivalencia, que relaciona el interior y la clausura de un conjunto, se deduce fácilmente

$$\overline{\mathcal{H} \setminus Z} = \mathcal{H} \setminus Z^\circ.$$

A partir de estas definiciones, probaremos dos propiedades conocidas de espacios métricos que usaremos en el Capítulo 6.

Proposición A.1.1. *Sea $\mathcal{H}$ un espacio de Hilbert y $\mathcal{S} \subseteq \mathcal{H}$ un subespacio, entonces:*

1. *Si $\mathcal{S}^\circ \neq \emptyset$, entonces $\mathcal{H} = \mathcal{S}$,*

2. *Si $\mathcal{S}^\circ = \emptyset$, entonces $\mathcal{H} \setminus \mathcal{S}$ es denso en $\mathcal{H}$.*

Demostración. 1. Supongamos que $\mathcal{S}^\circ \neq \emptyset$, entonces existen $x \in \mathcal{S}$ y $r > 0$ tal que $B(x,r) \subseteq \mathcal{S}$. Sea $z \in \mathcal{H}$, consideremos $y = \frac{x}{2} + \frac{r}{2\|z\|}z$, entonces $y \in B(x,r) \subseteq \mathcal{S}$. Como $\mathcal{S}$ es un subespacio, entonces $z = (y - \frac{x}{2})\frac{2\|z\|}{r} \in \mathcal{S}$ y entonces $\mathcal{H} = \mathcal{S}$.
2. Si $\mathcal{S}^\circ = \emptyset$, entonces $\overline{\mathcal{H} \setminus \mathcal{S}} = \mathcal{H} \setminus \mathcal{S}^\circ = \mathcal{H}$. □

Bases ortonormales en los espacios de Hilbert

En un espacio de Hilbert, todo sistema ortonormal $\{\psi_k : k \in \mathbb{N}\}$ cumple la desigualdad de Bessel, es decir cumple que si $x \in \mathcal{H}$, entonces

$$\sum_{k\geq 1} |\langle x, \psi_k \rangle|^2 \leq \|x\|^2.$$

Por otra parte, todo espacio de Hilbert separable $\mathcal{H}$, tiene una base ortonormal numerable. Llamaremos base ortonormal de un espacio de Hilbert, a todo conjunto $A \subset \mathcal{H}$ ortonormal maximal, es decir que si $A \subset B \subset H$ y B es un conjunto ortonormal entonces $B = A$. El siguiente resultado es bien conocido y es válido para espacios de Hilbert arbitrarios. Sin embargo, para simplificar supondremos que $\mathcal{H}$ es un espacio de Hilbert separable. Entonces si $\{\psi_k : k \in \mathbb{N}\}$, es un conjunto ortonormal de $\mathcal{H}$ son equivalentes:

1. Si $x \in \mathcal{H}$ verifica $\langle x, \psi_k \rangle = 0$ para todo $k \in \mathbb{N}$ entonces $x = 0$,

2. $\{\psi_k : k \in \mathbb{N}\}$ es una base ortonormal de $\mathcal{H}$,

3. si $x \in \mathcal{H}$ entonces $x = \sum_{k\geq 1} \langle x, \psi_k \rangle \psi_k = \lim_{N\to\infty} \sum_{k=1}^N \langle x, \psi_k \rangle \psi_k$,

4. si $x, y \in \mathcal{H}$ entonces $\langle x, y \rangle = \sum_{k\geq 1} \langle x, \psi_k \rangle \langle \psi_k, y \rangle$, (ecuación de Parseval),

5. si $x \in \mathcal{H}$ entonces $\|x\|^2 = \sum_{k\geq 1} |\langle x, \psi_k \rangle|^2$, (igualdad de Parseval).

Operador adjunto

Dado $T \in L(\mathcal{H}, \mathcal{K})$, el único operador $S \in L(\mathcal{H}, \mathcal{K})$ tal que

$$\langle Th, k \rangle = \langle h, Sk \rangle \text{ para todo } h \in \mathcal{H},\ k \in \mathcal{K},$$

se llama operador adjunto de T y se denota como $S = T^*$. La existencia y unicidad de dicho operador está garantizada por el teorema de Representación de Riesz [19, Capítulo I, Teorema 3.4].

Sea $T \in L(\mathcal{H})$. Entonces,

1. T es positivo si $\langle Tx, x \rangle \geq 0$ para todo $x \in \mathcal{H}$,
2. T es hermitiano ó autoadjunto si $T = T^*$,
3. T es normal si $TT^* = T^*T$,
4. T es unitario, si T es inversible y $T^{-1} = T^*$.

Se puede probar que si T es positivo, entonces T es hermitiano. Se deduce inmediatamente que si T es autoadjunto o unitario, entonces es normal.

A.1.1 *Teoría espectral*

Dado un operador $T \in L(\mathcal{H})$, el conjunto resolvente de T se notará por $\rho(T)$ y se define como

$$\rho(T) = \{\lambda \in \mathbb{C} : \text{ existe } (\lambda I - T)^{-1}\},$$

el complemento de dicho conjunto, se llamara espectro de T y se notará como $\sigma(T) = \rho(T)^C$. El conjunto $\sigma_P(T)$, es el espectro puntual de T y se define como

$$\sigma_P(T) = \{\lambda \in \mathbb{C} : N(\lambda I - T) \neq \{0\}\};$$

los elementos de $\sigma_P(T)$ se llamarán los autovalores de T.

Los siguientes teoremas, probados en [75], serán usados ampliamente en este trabajo. El primero de los teoremas que veremos se llama comúnmente Teorema Espectral.

Teorema A.1.2 (Teorema espectral)**.** *Sea $T \in L(\mathcal{H})$ y r un polinomio con coeficientes complejos, entonces*

$$\sigma(r(T)) = \{r(\lambda) : \lambda \in \sigma(T)\}.$$

Dado, un conjunto compacto K, denotaremos por $\mathcal{C}(K)$ el conjunto de funciones continuas de K a $\mathbb{R}$.

Usando el Teorema de Stone-Weirstrass [19, Capítulo I,Teorema 5.6] y el Teorema Espectral, y teniendo en cuenta que el espectro de un operador autoadjunto acotado, es un conjunto compacto de la recta real [75, Teorema 1.5.1] y [75, Teorema 1.7.6], es posible demostrar el siguiente resultado, conocido como Teorema de Cálculo Fucional [75, Teorema 1.7.7]. y

Teorema A.1.3 (Teorema de cálculo funcional)**.** *Si $T = T^* \in L(\mathcal{H})$, existe una única función lineal $f \to f(T)$ de $\mathcal{C}(\sigma(T))$ a $L(\mathcal{H})$ tal que:*

1. *$f(T)$ tiene su significado esperado cuando f es un polinomio,*
2. $\|f(T)\| = \|f\|_\infty = sup\{|f(x)| : x \in \sigma(T)\}, \;\; f \in \mathcal{C}(\sigma(T))$.

   Es más, para todo $f, g \in \mathcal{C}(\sigma(T))$, vale que
3. $(fg)(T) = f(T)g(T)$,

4. $\overline{f}(T) = (f(T))^*$,

5. $f(T)$ *es normal,*

6. $f(T)S = Sf(T)$, *si* $S \in L(\mathcal{H})$ *y* $TS = ST$,

7. *si* $\lambda \in \sigma_P(T)$, $x \in \mathcal{H}$ *y* $Tx = \lambda x$, *entonces*

$$f(T)x = f(\lambda)x.$$

A partir del Teorema de Cálculo Funcional, es posible probar para un operador T positivo, la existencia de un único operador B positivo, tal que $T = B^2$, dicho operador se llamará la raíz cuadrada de T y lo probaremos en el siguiente teorema.

Teorema A.1.4. *Sea* $T \in L(\mathcal{H})$ *autoadjunto, entonces las siguientes condiciones son equivalentes:*

1. $\sigma(T)$ *consiste de números reales no negativos,*

2. $T = B^2$, *para algún operador positivo* $B \in L(\mathcal{H})$,

3. $T = A^*A$, *para algún* $A \in L(\mathcal{H})$,

4. $T \geq 0$.

Cuando esta condiciones se satisfacen, el operador B *de* 2. *es único.*

Demostración. Es claro que 2. implica 3. Si $T = A^*A$, entonces $\langle Tx, x \rangle = \langle Ax, Ax \rangle \geq 0$, para cada $x \in \mathcal{H}$, entonces 3. implica 4. Si 1. se satisface, la función a valores reales, $g(\lambda) = \lambda^{1/4}$, es continua en $\sigma(T)$ y $[g(\lambda)]^4 = \lambda$. Entonces, por el Teorema A.1.3, $S = g(T)$ cumple $S^4 = T$. Llamando $B = S^2$, entonces B es un operador positivo y $B^2 = T$, por lo tanto 1. implica 2.

Supongamos que $T \geq 0$, y definamos h en $\mathcal{C}(\sigma(T))$ por

$$h(\lambda) = \begin{cases} 0 & \text{si } \lambda \geq 0 \\ (-\lambda)^{1/2} & \text{si } \lambda < 0. \end{cases}$$

Como h es una función a valores reales y $h(\lambda)\lambda h(\lambda) = -[h(\lambda)]^4$ cuando $\lambda \in \sigma(T)$, el operador $S = h(T)$ es autoadjunto y $STS = -S^4$. Como 3. implica 4., vale que $S^4 \geq 0$, entonces

$$0 \geq -S^4 = STS = STS^* \geq 0,$$

y por lo tanto $0 = S^4 = [h(\lambda)]^4$, como el mapa $f \to f(T)$ es inyectivo, se sigue que h^4 es el elemento nulo de $\mathcal{C}(\sigma(T))$. Esto, junto con la definición de h, muestra que $\sigma(T)$, consiste de números reales no negativos. Entonces 4. implica 1. Y las cuatro equivalencias del teorema quedan probadas.

Solo queda por probar que si $T \geq 0$ entonces el operador B de 3. es único. Sea $B = f(T)$, donde $f \in \mathcal{C}(\sigma(T))$, definido por $f(\lambda) = \lambda^{1/2}$. Si $\{r_n\}_{n\geq 1}$ es una sucesión de polimios que converge uniformemente en $\sigma(T)$ a f (Teorema de Weierstrass), entonces

$$\|B - r_n(T)\| \underset{n\to\infty}{\to} 0.$$

Sea D otro operador positivo tal que $D^2 = T$. Dado que $\sigma(D)$ consiste en números reales no negativos, y $\sigma(T) = \{\mu^2 : \mu \in \sigma(D)\}$, el polinomio $q_n(\mu) = r_n(\mu^2)$ converge uniformemente en $\sigma(D)$, a la función $f(\mu^2) = \mu = v(\mu)$, donde v es la función identidad en $\sigma(D)$. Entonces

$$\|D - r_n(T)\| = \|D - q_n(D)\| = \|v - q_n\|_\infty \underset{n\to\infty}{\to} 0.$$

Entonces

$$\|B - D\| \leq \|B - r_n(T)\| + \|D - r_n(T)\| \underset{n\to\infty}{\to} 0,$$

y tenemos que $B = D$. □

Definición. Sea $T \in L(\mathcal{H})$. Como T^*T es autoadjunto y positivo, por el teorema anterior, podemos definir

$$|T| = (T^*T)^{1/2},$$

siendo $B = |T|$, el único operador positivo tal que $B^2 = T^*T$.

A continuación probaremos el Teorema de Descomposición Polar, para probar dicho teorema es necesario enunciar un lema técnico de fácil demostración.

Lema A.1.5. *Sea $\mathcal{H}$ un espacio de Hilbert, $T \in L(\mathcal{H})$ entonces:*

1. $\|Tx\| = \|\, |T|\, x\|$ *y* $N(T) = N(|T|)$,
2. *R(T) es cerrado si y sólo si $R(|T|)$ es cerrado.*

Demostración. 1. $\|Tx\|^2 = \langle Tx, Tx \rangle = \langle T^*Tx, x \rangle = \langle |T|^2 x, x \rangle = \langle |T|x, |T|x \rangle = \|\, |T|\, x\|^2$, y a partir de esto es inmediato que $N(T) = N(|T|)$.

2. Por la Proposición 2.1.11, si $R(T)$ es cerrado, vale que T es acotado inferiormente, es decir, existe $\alpha > 0$ tal que

$$\|Tx\| \geq \alpha\|x\|, \text{ para todo } x \in R(T)^\perp,$$

usando el ítem 1., tenemos que

$$\|\, |T|\, x\| = \|Tx\| \geq \alpha\|x\|, \text{ para todo } x \in R(T)^\perp = R(|T|)^\perp,$$

por lo tanto, nuevamente por la Proposición 2.1.11, $R(|T|)$ es cerrado. La recíproca se prueba de igual manera.

□

Definición. Sean $\mathcal{H}$ un espacio de HIlbert y $U \in L(\mathcal{H})$, diremos que U es una *isometría parcial* si

$$\|Ux\| = \|x\|, \text{ para cada } x \in N(U)^{\perp}.$$

Teorema A.1.6 (Descomposición polar)**.** . *Sean $\mathcal{H}$ un espacio de HIlbert y $T \in L(\mathcal{H})$. Entonces existe una única isometría parcial $U \in L(\mathcal{H})$, tal que*

$$T = U|T| \text{ y } N(U) = N(T).$$

Demostración. Definimos $U : R(|T|) \to \mathcal{H}$ como

$$Uy = Tx,$$

donde $y = |T|x$. Entonces U está bien definido por el ítem 1. del lema anterior, es líneal y además

$$\|Uy\| = \|Tx\| = \| \, |T|x\| = \|y\|.$$

Por lo tanto U resulta una isometría, en particular es continua. Entonces podemos extenderla de manera isométrica a $\overline{R(|T|)}$. Ahora bien, sabiendo que $|T|^* = |T|$ y usando el ítem 1. del lema anterior tenemos:

$$\overline{R(|T|)} = N(|T|)^{\perp} = N(T)^{\perp}.$$

En definitiva, definimos U sobre $N(T)^{\perp}$. Luego extendemos U sobre $N(T)$ como el operador nulo. Como $\mathcal{H} = \overline{R(|T|)} \oplus N(|T|)^{\perp} = N(T) \oplus N(T)^{\perp}$, extendemos U por linealidad a todo el espacio y está bien definida.

Sea $x \in \mathcal{H}$, $x = a + b$, con $a \in N(T)$, $b \in N(T)^{\perp}$. Entonces

$\|Ux\|^2 = \|U(a+b)\|^2 = \|Ub\|^2 = \|b\|^2 \leq \|a\|^2 + \|b\|^2 = \|x\|^2$, y por lo tanto $\|U\| = 1$.

Veamos que $N(U) = N(T)$. Puesto que U es el operador nulo sobre $N(T)$, es claro que $N(T) \subseteq N(U)$. Sea $x \in \mathcal{H}$, $x = a + b$, con $a \in N(T)$, $b \in N(T)^{\perp}$. Entonces

$$0 = Ux = Ua + Ub = Ub,$$

pues $U|_{N(T)} \equiv 0$, por lo tanto

$$0 = \|Ux\| = \|Ub\| = \|b\|,$$

porque U es isometría sobre $N(T)^{\perp}$. Es decir, $x = a \in N(T)$.

La unicidad es clara del hecho que $\mathcal{H}$ es suma directa de $R(|T|)$ y $N(T)$, y dos tales U, V tienen por núcleo a $N(T)$ y valen lo mismo restrigidos a $R(|T|)$. $\square$

Definición. Sea $A \in L(\mathcal{H})$, se define la parte real e imaginaria de A como:

$$Re(A) = \frac{A + A^*}{2} \text{ e } Im(A) = \frac{A - A^*}{2i}.$$

Observar que $A = Re(A) + i \; Im(A)$, y que tanto $Re(A)$ como $Im(A)$ son operadores autoadjuntos.

Lema A.1.7. *Cada elemento $T \in L(\mathcal{H})$ es una combinación lineal finita de operadores unitarios.*

Demostración. Dado que cada elemento $T \in L(\mathcal{H})$ se puede representar en términos de su parte real e imaginaria, las cuales son autoadjuntas, podemos considerar el caso en que $T = T^*$ y $\|T\| \leq 1$. En este caso, el espectro $\sigma(T) \subseteq [-1, 1]$, [75, Teorema 1.5.1], entonces podemos definir una función continua f en $\sigma(T)$, por

$$f(t) = t + i\sqrt{1 - t^2}.$$

Sea $U = f(T)$, el operador correspondeintes a f en el cálculo funcional descrita en A.1.3. Entonces, dado que

$$t = \frac{f(t) + \overline{f(t)}}{2} \quad \text{y} \quad f(t)\overline{f(t)} = \overline{f(t)}f(t) = 1,$$

se sigue del Teorema A.1.3, que

$$T = \frac{U + U^*}{2} \quad \text{y} \quad UU^* = U^*U = I.$$

□

A.2 OPERADORES COMPACTOS

Definición. Una transformación lineal $T : \mathcal{H} \to \mathcal{H}$ es compacta si $T(B[0, 1])$ tiene clausura compacta, donde $B[0, 1] = \{x \in \mathcal{H} : \|x\| \leq 1\}$, es la bola unitaria de $\mathcal{H}$. El conjunto de operadores compactos de $\mathcal{H}$ en $\mathcal{H}$ se denotará $K(\mathcal{H})$.

Observar que $K(\mathcal{H}) \subset L(\mathcal{H})$, que $K(\mathcal{H})$ es un espacio vectorial cerrado y además es un ideal bilátero, es decir que si $A \in L(\mathcal{H}), B \in L(\mathcal{H})$ y $T \in K(\mathcal{H})$, entonces TA y $BT \in K(\mathcal{H})$.

Diremos que un operador T se llama de rango finito si $R(T)$ tiene dimensión finita. El conjunto de operadores de rango finito de $\mathcal{H}$ en $\mathcal{H}$ forma un espacio vectorial contenido en $K(\mathcal{H})$. Además, si T es de rango finito entonces T^* es de rango finito y las dimensiones de $R(T)$ y $R(T^*)$ coinciden.

El siguiente conocido teorema, cuya prueba se encuentra en [75, Teorema 1.8.7], permite ver que el conjunto de operadores de rango finito, al que notaremos $\mathcal{F}$ es denso en $K(\mathcal{H})$.

Teorema A.2.1. *Si $T \in L(\mathcal{H})$, son equivalentes:*

1. *T es compacto,*

2. *T^* es compacto,*

3. *existe una sucesión* $\{T_n\}_{n\geq 1}$ *de operadores en* $\mathcal{F}$ *tal que*

$$\|T - T_n\| \to 0.$$

Por último enunciaremos el importante teorema de Diagonalización para operadores compactos normales. Ver prueba en [75, Teorema 1.9.2].

Teorema A.2.2 (Diagonalización de operadores compactos normales)**.** *Sea* $T \in K(\mathcal{H})$ *un operador normal, sea* $\{\lambda_n\}_{n\geq 1}$ *la sucesión de autovalores de* T *(ordenados decrecientemente de acuerdo a su valor absoluto y contados de acuerdo a sus multiplicidades), y sea* $\{\phi_n\}_{n\geq 1}$ *la sucesión de autovectores ortonormales correspondiente. Entonces*

$$Tx = \sum_{n\geq 1} \lambda_n \langle x, \phi_n \rangle \phi_n.$$

Usando el teorema de descomposición polar, podemos probar que si $T \in K(\mathcal{H})$ (no necesariamente normal) entonces, existe una sucesión decreciente de números reales positivos $\{\mu_n\}_{n\geq 1}$ y $\{\psi_n\}_{n\geq 1}$, $\{\phi_n\}_{n\geq 1}$ sucesiones ortonormales en $\mathcal{H}$. Tal que

$$Tx = \sum_{n\geq 1} \mu_n \langle x, \phi_n \rangle \psi_n. \tag{A.1}$$

De hecho, sea $T = U|T|$ la descomposición polar de T, entonces por el teorema de cálculo funcional, $|T|$ es un operador positivo y compacto. Sea $\{\mu_n\}_{n\geq 1}$, la sucesión de autovalores no nulos de $|T|$ (decreciente de números reales positivos contados de acuerdo a sus multiplicidades), por el Teorema A.2.2 existe $\{\phi_n\}_{n\geq 1}$ una sucesión ortonormal en $\mathcal{H}$ tal que

$$|T|x = \sum_{n\geq 1} \mu_n \langle x, \phi_n \rangle \phi_n. \tag{A.2}$$

Sea $\psi_k = U\phi_k$. Como U es una isometría parcial en $N(T)^\perp$ y $\phi_n = \mu_n^{-1}|T|\phi_n \in R(|T|) \subseteq N(T)^\perp$, entonces $U^*U\phi_n = \phi_n$, por lo tanto

$$\langle \psi_m, \psi_n \rangle = \langle U\phi_m, U\phi_n \rangle = \langle U^*U\phi_m, \phi_n \rangle = \langle \phi_m, \phi_n \rangle,$$

entonces $\{\psi_n\}_{n\geq 1}$ es un conjunto ortonormal en $\mathcal{H}$. Entonces, volviendo a la ecuación A.2, tenemos

$$Tx = U|T|x = \sum_{n\geq 1} \mu_n \langle x, \phi_n \rangle U\phi_n = \sum_{n\geq 1} \mu_n \langle x, \phi_n \rangle \psi_n.$$

Finalizamos la sección con una proposición que usa varios de los resultados presentados hasta ahora y que se utilizará frecuentemente a lo largo de este trabajo.

Proposición A.2.3. *Sea* $T \in K(\mathcal{H})$ *un operador positivo y sean* $r, s > 0$ *entonces*

$$N(T^r) = N(T^s).$$

Demostración. Supongamos que $T \in K(\mathcal{H})$ y que T es un operador positivo. Sea $\{\lambda_n\}_{n\geq 1}$, la sucesión de autovalores reales no nulos (ver Teorema A.1.4) de T (sucesión decreciente de números reales positivos y contados de acuerdo a sus multiplicidades), y sea $\{\phi_n\}_{n\geq 1}$ la sucesión de autovectores ortonormales correspondiente. Como $T = T^*$, por el Teorema 6.1.13, vale que

$$Tx = \sum_{n\in\mathcal{N}_1} \lambda_n \langle x, \phi_n \rangle \phi_n,$$

donde $\mathcal{N}_1 = \{n \in \mathbb{N} : \lambda_n > 0\}$.

Por el Teorema de Cálculo funcional A.1.3, vale que si $r > 0$ entonces

$$T^r x = \sum_{n\in\mathcal{N}_1} \lambda_n^r \langle x, \phi_n \rangle \phi_n,$$

donde se usó que si $\lambda_n > 0$, entonces $\lambda_n^r > 0$. Sea $x \in N(T^r)$, entonces

$$0 = \|T^r x\|^2 = \sum_{n\in\mathcal{N}_1} \lambda_n^r |\langle x, \phi_n \rangle|^2,$$

entonces como $\lambda_n^r > 0$ para todo $n \in \mathcal{N}_1$ vale que $\langle x, \phi_n \rangle = 0$, para todo $n \in \mathcal{N}_1$. Entonces

$$T^s x = \sum_{n\in\mathcal{N}_1} \lambda_n^s \langle x, \phi_n \rangle \phi_n = 0,$$

y $x \in N(T^s)$, concluyendo que

$$N(T^r) \subseteq N(T^s).$$

La otra inclusión se prueba de manera similar.

□

B

PRUEBA DE LAS PROPIEDADES ENUNCIADAS DE LOS ESPACIOS S_p

En este anexo se realizan las pruebas a las propiedades enunciadas de los espacios de Schatten S_p que se encuentran en el Capítulo 3.

B.1 OPERADORES EN LA CLASE DE SCHATTEN

Lema B.1.1. *Supongamos $1 \leq p < \infty$, T es un operador compacto y autoadjunto en $\mathcal{H}$, y $\{\lambda_n\}_{n\geq 1}$ es la sucesión de autovalores no nulos de T, contados de acuerdo a su multiplicidad. Entonces:*

1. *Si $T \in S_p$ entonces $\sum_{n\geq 1} |\lambda_n|^p < \infty$,*
2. *si $\sum_{n\geq 1} |\lambda_n|^p < \infty$, entonces $T \in S_p$ y para cualquier sistema ortonormal $\{\psi_k : k \in \mathbb{N}\}$ en $\mathcal{H}$,*

$$\sum_{k\geq 1} |\langle T\psi_k, \psi_k \rangle|^p \leq \sum_{n\geq 1} |\lambda_n|^p.$$

Demostración. Por el Teorema A.2.2, existe una sucesión ortonormal $\{\phi_n\}_{n\geq 1} \subseteq \mathcal{H}$ tal que

$$Tx = \sum_{n\geq 1} \lambda_n \langle x, \phi_n \rangle \phi_n \text{ para todo } x \in \mathcal{H}.$$

Si $T \in S_p$ entonces por definición

$$\sum_{n\geq 1} |\lambda_n|^p = \sum_{n\geq 1} |\langle T\phi_n, \phi_n \rangle|^p < \infty.$$

Recíprocamente, supongamos que $\sum_{n\geq 1} |\lambda_n|^p < \infty$, entonces si $\{\psi_k : k \in \mathbb{N}\}$ es cualquier sistema ortonormal en $\mathcal{H}$, entonces

$$\langle T\psi_k, \psi_k \rangle = \sum_{n\geq 1} \lambda_n |\langle \psi_k, \phi_n \rangle|^2.$$

Si q es el conjugado de p, es decir $\frac{1}{q} + \frac{1}{p} = 1$, usando la desigualdad de Hölder [75, Lema 1.3.2] (con la interpretación usual cuando $p = 1$ y $q = \infty$), tenemos:

$$|\langle T\psi_k, \psi_k \rangle| \leq \sum_{n\geq 1} |\lambda_n| |\langle \psi_k, \phi_n \rangle|^{\frac{2}{p}} |\langle \psi_k, \phi_n \rangle|^{\frac{2}{q}} \leq$$

$$\Big(\sum_{n\geq 1} |\lambda_n|^p |\langle \psi_k, \phi_n \rangle|^2\Big)^{\frac{1}{p}} \Big(\sum_{n\geq 1} |\langle \psi_k, \phi_n \rangle|^2\Big)^{\frac{1}{q}}.$$

Ya que $\sum_{n\geq 1} |\langle \psi_k, \phi_n \rangle|^2 \leq \|\psi_k\|^2 = 1$, tenemos

$$\sum_{k\geq 1} |\langle T\psi_k, \psi_k \rangle|^p \leq \sum_{k\geq 1}\sum_{n\geq 1} |\lambda_n|^p |\langle \psi_k, \phi_n \rangle|^2 \leq$$

$$\sum_{n\geq 1} |\lambda_n|^p \sum_{k\geq 1} |\langle \phi_n, \psi_k \rangle|^2 \leq \sum_{n\geq 1} |\lambda_n|^p \|\phi_n\|^2 = \sum_{n\geq 1} |\lambda_n|^p,$$

donde en la ultima desigualdad usamos la desigualdad de Bessel. Entonces $T \in S_p$ y 2. se satisface. □

Es bien sabido que un operador T en $\mathcal{H}$ es compacto si y sólo si es el límite en norma de operadores de rango finito, ver Teorema A.2.1. El próximo resultado muestra que pertencer a S_p también es equivalente a ser límite en norma de operadores de rango finito.

Teorema B.1.2. *Sea $T \in L(\mathcal{H})$ y $1 \leq p < \infty$. Entonces $T \in S_p$ si y sólo si existe una sucesión $\{F_n\}_{n\geq 1}$ de operadores en $\mathcal{H}$, tal que F_n tiene rango finito, no mayor que n y*

$$\sum_{n\geq 1} \|T - F_n\|^p < \infty.$$

Demostración. Sea $\mathcal{D}$ la clase de operadores T en $\mathcal{H}$ para los cuales dicha sucesión $\{F_n\}_{n\geq 1}$ existe. Dado $S, T \in \mathcal{D}$ y escalares α, β, elijamos F_n y G_n operadores de rango finito no mayor a n para los cuales se cumple que $\sum_{n\geq 1} \|S - F_n\|^p < \infty$ y $\sum_{n\geq 1} \|T - G_n\|^p < \infty$, y definamos la sucesión $\{H_n\}_{n\geq 1}$ por

$$H_1 = 0, \quad H_{2n} = H_{2n+1} = \alpha F_n + \beta G_n.$$

Entonces H_n tiene rango finito no mayor a n, y

$$(\sum_{n\geq 2} \|\alpha S + \beta T - H_n\|^p)^{1/p} =$$

$$= 2^{1/p}(\sum_{n\geq 1} \|\alpha(S - F_n) + \beta(T - G_n)\|^p)^{1/p} \leq$$

$$2^{1/p}(\sum_{n\geq 1} (|\alpha|\|S - F_n\|)^p)^{1/p} + 2^{1/p}(\sum_{n\geq 1} (|\beta|\|T - G_n\|)^p)^{1/p} < \infty,$$

donde se usó la desigualdad de Minkowski [75, Lema 1.3.1], entonces $\alpha S + \beta T \in \mathcal{D}$. Es más $S^* \in \mathcal{D}$, ya que F_n^* tiene rango finito no mayor a n y

$$\sum_{n\geq 1} \|S^* - F_n^*\|^p = \sum_{n\geq 1} \|S - F_n\|^p < \infty.$$

Finalmente, S es compacto ya que es el límite en norma de la sucesión de operadores de rango finito $\{F_n\}_{n\geq 1}$.

Estos resultados, prueban que $\mathcal{D}$ es subespacio lineal de $L(\mathcal{H})$ que consiste de operadores compactos que también contienen sus adjuntos. Vimos que S_p también tiene esas propiedades. Para probar que $\mathcal{D} = S_p$ es suficiente con probar que un operador compacto y autoadjunto $T \in \mathcal{D}$ si y sólo si $T \in S_p$, (pensar en la descomposición en parte real e imaginaria de un operador $S \in L(\mathcal{H})$, definida en la sección A.1.1).

Dado dicho T, por el Teorema A.2.2, podemos suponer que

$$Tx = \sum_{m\geq 1} \lambda_m \langle x, \phi_m \rangle \phi_m,$$

donde $\{\phi_m\}_{m\geq 1}$ es una sucesión ortonormal y $\{\lambda_m\}_{m\geq 1}$ es la sucesión de autovalores no nulos de T (ordenados por módulo decreciente, contados de acuerdo a su mulplicidad y seguido por una sucesión infinita de ceros si T es de rango finito).

Si $T \in S_p$, se define el operador F_n por

$$F_n x = \sum_{m=1}^{n} \lambda_m \langle x, \phi_m \rangle \phi_m,$$

entonces F_n tiene rango finito no mayor que n,

$$(T - F_n)x = \sum_{m=n+1}^{\infty} \lambda_m \langle x, \phi_m \rangle \phi_m,$$

entonces por [75, Lema 1.9.1], tenemos que

$$\|T - F_n\| = sup\{|\lambda_m| \colon m > n\} = |\lambda_{n+1}|.$$

Por el Lema 3.1.1

$$\sum_{n\geq 1} \|T - F_n\|^p = \sum_{n\geq 1} |\lambda_{n+1}|^p < \infty,$$

entonces $T \in \mathcal{D}$.

Recíprocamente, supongamos que $T \in \mathcal{D}$. Sea G_n el operador de rango finito no mayor que $n \in \mathbb{N}$ para el cual $\sum_{n\geq 1} \|T - G_n\|^p < \infty$. Para un n fijo, podemos elegir vector $x_m, y_m \in \mathcal{H},\ 1 \leq m \leq n$ tal que

$$G_n x = \sum_{m=1}^{n} \langle x, x_m \rangle y_m.$$

Entonces $G_n(\alpha_1\phi_1 + \cdots + \alpha_{n+1}\phi_{n+1}) = 0$ siempre que los escalares $\alpha_1, \cdots, \alpha_{n+1}$ satisfagan las n ecuaciones lineales $\langle \alpha_1\phi_1 + \cdots + \alpha_{n+1}\phi_{n+1}, x_m \rangle = 0$, $m = 1, \cdots, n$.

Con $\alpha_1, \cdots, \alpha_{n+1}$ (no todos nulos) satisfaciendo estas ecuaciones tenemos

$$\|(T - G_n)(\alpha_1\phi_1 + \cdots + \alpha_{n+1}\phi_{n+1})\| = \|T(\alpha_1\phi_1 + \cdots + \alpha_{n+1}\phi_{n+1})\| =$$

$$= \|\lambda_1\alpha_1\phi_1 + \cdots + \lambda_{n+1}\alpha_{n+1}\phi_{n+1}\| = (|\lambda_1\alpha_1|^2 + \cdots + |\lambda_{n+1}\alpha_{n+1}|^2)^{\frac{1}{2}} \geq$$

$$|\lambda_{n+1}|(|\alpha_1|^2 + \cdots + |\alpha_{n+1}|^2)^{1/2} = |\lambda_{n+1}|\|\alpha_1\phi_1 + \cdots + \alpha_{n+1}\phi_{n+1}\|.$$

Por lo tanto $|\lambda_{n+1}| \leq \|T - G_n\|$, y entonces

$$\sum_{n\geq 1} |\lambda|^p \leq |\lambda_1|^p + \sum_{n\geq 1} \|T - G_n\|^p < \infty,$$

y por el Lema 3.1.1, $T \in S_p$. $\square$

B.1.1 *El espacio de Banach S_p*

Lema B.1.3. *Sea $1 \leq q \leq p < \infty$. Entonces las siguientes condiciones son equivalentes:*

1. $T \in S_p$,
2. $|T| \in S_p$,
3. $|T|^{p/q} \in S_q$.

Demostración. Como S_p es un ideal en $L(\mathcal{H})$ las equivalencias entre 1. y 2., se siguen de las ecuaciones, $T = U|T|$ y $|T| = U^*T$.

Por otra parte, si 1. ó 2. se satisfacen, entonces $|T| = (|T|^{p/q})^{q/p}$, es compacto por el Teorema 3.1.5, y podemos suponer que $|T|$ está representado por la ecuación 3.1. Por el Lema 3.1.1, $|T| \in S_p$ si y sólo si $\sum_{n\geq 1} \mu_n^p < \infty$, esto es equivalente a que $\sum_{n\geq 1}(\mu_n^{p/q})^q < \infty$, lo que nuevamente por el Lema 3.1.1, es una condición suficiente y necesaria para que $|T|^{p/q} \in S_q$, entonces 2. y 3. son equivalentes. □

Definición. Sea $\mathcal{F}$ el conjunto de todos los operadores de rango finito de $\mathcal{H}$ en $\mathcal{H}$. Sean $x, y \in \mathcal{H}$, definimos $x \otimes y \in \mathcal{F}$ como

$$(x \otimes y)z = \langle z, y \rangle x, \qquad z \in \mathcal{H}.$$

Observar que si $\{\phi_j : j \in \mathbb{N}\}$ es una base ortonormal de $\mathcal{H}$, entonces

$$tr(x \otimes y) = \sum_{j\geq 1} \langle (x \otimes y)\phi_j, \phi_j \rangle = \sum_{j\geq 1} \langle \phi_j, y \rangle \langle x, \phi_j \rangle = \langle x, y \rangle,$$

donde usamos la ecuación de Parseval.

Observar además que, dados $u, v \in \mathcal{H}$, tenemos que

$$\langle (x \otimes y)u, v \rangle = \langle \langle u, y \rangle x, v \rangle = \langle u, y \rangle \langle x, v \rangle = \left\langle u, \overline{\langle x, v \rangle} y \right\rangle =$$

$$= \langle u, \langle v, x \rangle y \rangle = \langle u, (y \otimes x)v \rangle,$$

por lo tanto $(x \otimes y)^* = y \otimes x$.

Por último, veamos que

$$\|x \otimes y\|_p = \|x\| \, \|y\|. \tag{B.1}$$

Para ello es suficiente considerar el caso en que $\|x\| = \|y\| = 1$. En este caso, sea $x \otimes y = U|x \otimes y|$, la descomposición polar de $x \otimes y$.

Veamos que $|x \otimes y| = ((x \otimes y)^*(x \otimes y))^{1/2} = y \otimes y$. De hecho, dado $z \in \mathcal{H}$, tenemos que

$$(x \otimes y)^*(x \otimes y)z = (y \otimes x) \langle z, y \rangle x = \langle z, y \rangle \langle x, x \rangle y = \langle z, y \rangle y = (y \otimes y)z,$$

entonces $(x \otimes y)^*(x \otimes y) = (y \otimes y)$ y vale que

$$|x \otimes y| = ((x \otimes y)^*(x \otimes y))^{1/2} = (y \otimes y)^{1/2} = y \otimes y.$$

Entonces el único autovalor de $|x \otimes y|$ es el 1, y su multiplicidad es 1. Luego por (3.4), vale que

$$\|x \otimes y\|_p = 1.$$

A continuación enuciaremos un lema técnico, cuya prueba se encuentra en [75].

Lema B.1.4. *Sea $1 \leq p < \infty$, q el índice conjugado de p. Entonces $T \in S_p$ si y sólo si*

$$sup\{|tr\ (FT)| : F \in \mathcal{F}, \|F\|_q \leq 1\} < \infty,$$

donde $\mathcal{F}$ es el conjunto de los operadores de rango finito en $\mathcal{H}$. Cuando esto es así, el valor del supremo es $\|T\|_p$.

Teorema B.1.5. *Sea $1 \leq p < \infty$. Entonces $\|T\|_p$ es una norma en S_p, y con esta norma S_p es un espacio de Banach. Además, el conjunto $\mathcal{F}$ de los operadores de rango finito en $\mathcal{H}$ es un subespacio denso de S_p.*

Demostración. Dado que $tr(T)$ es un funcional lineal en S_1, y por el Lema B.1.4 vale que para todo $T \in S_p$

$$\|T\|_p = sup\{|tr\ (FT)| : F \in \mathcal{F}, \|F\|_q \leq 1\},$$

entonces se sigue que

$$\|T\|_p \geq 0, \quad \|\alpha T\|_p = |\alpha|\|T\|_p, \quad \|S + T\|_p \leq \|T\|_p + \|S\|_p,$$

para todo $S, T \in S_p$ y todo $\alpha \in \mathbb{C}$.

Por otra parte, sea $\{T_n\}_{n\geq 1}$ una sucesión de Cauchy en S_p. Entonces, dado $\varepsilon > 0$, existe $N(\varepsilon) \in \mathbb{N}$ tal que

$$\|T_m - T_n\|_p < \varepsilon, \text{ para todo } n, m \geq N(\varepsilon). \tag{B.2}$$

Entonces, por la ecuación (3.5), $\|T_m - T_n\| < \varepsilon$, para todo $n, m \geq N(\varepsilon)$, por lo tanto $\{T_n\}_{n\geq 1}$ es una sucesión de Cauchy en $L(\mathcal{H})$, y entonces converge en la norma de operadores a algún $T \in L(\mathcal{H})$. Si $x, y \in \mathcal{H}$ entonces

$$tr((x \otimes y)T_n) = tr(x \otimes T_n^* y) = \langle T_n x, y \rangle \underset{n\to\infty}{\to} \langle Tx, y \rangle = tr((x \otimes y)T).$$

Se sigue de la linealidad de tr, que $tr(FT_n) \underset{n\to\infty}{\to} tr(FT)$, para todo $F \in \mathcal{F}$.

Sea $F \in \mathcal{F}$ tal que $\|F\|_q \leq 1$, por el Lema B.1.4 y por (B.2) tenemos que

$$|tr(FT_m - FT_n)| < \varepsilon, \text{ para todo } n, m \geq N(\varepsilon),$$

y tomando el límite del lado derecho cuando $n \to \infty$ obtenemos

$$|tr(FT_m - FT)| \leq \varepsilon, \text{ para todo } m \geq N(\varepsilon).$$

Entonces

$$sup\{|tr\ (FT_m - FT)| : F \in \mathcal{F}, \|F\|_q \leq 1\} \leq \varepsilon,$$

cuando $m \geq N(\varepsilon)$, y el Lema B.1.4 muestra que

$$T_m - T \in S_p, \quad \|T_m - T\|_p \leq \varepsilon, \ (m \leq N(\varepsilon)).$$

Se sigue entonces, que $T \in S_p$ y que T_n converge en la norma de S_p a T. Esto prueba que S_p es completo, y por lo tanto un espacio de Banach.

Finalmente, veamos que cada $T \in S_p$ pertenece a $\overline{\mathcal{F}}$. Podemos suponer que T está representado por la ecuación (3.2). Para $m \in \mathbb{N}$, definamos $T_m \in \mathcal{F}$ por

$$T_m x = \sum_{n=1}^{m} \mu_n \langle x, \phi_n \rangle \psi_n.$$

Entonces

$$(T - T_m)x = \sum_{n=m+1}^{\infty} \mu_n \langle x, \phi_n \rangle \psi_n,$$

entonces

$$\|T - T_m\|_p = (\sum_{m+1}^{\infty} \mu_n^p)^{1/p} \underset{m\to\infty}{\to} 0.$$

Por lo tanto $T \in \overline{\mathcal{F}}$. □

B.1.2 *La clase de Hilbert-Schmidt*

Definición. El ideal S_2 en $L(\mathcal{H})$ es llamado la clase de operadores de Hilbert-Schmidt en $\mathcal{H}$.

Ya probamos que S_2 es un espacio de Banach con respecto a la norma $\|\cdot\|_2$. El siguiente teorema muestra que S_2 es, de hecho, un espacio de Hilbert y tiene una base ortonormal que consiste en operadores de la forma $x \otimes y$.

Teorema B.1.6. *La clase de Schmidt S_2 es un espacio de Hilbert con respecto al producto interno $\langle \cdot, \cdot \rangle$ definido por*

$$\langle S, T \rangle = tr(T^* S),$$

para $S, T \in S_2$. La norma derivada de este producto interno es $\|\cdot\|_2$. Por último si $\{\phi_j\}_{j\geq 1}$ y $\{\psi_j\}_{j\geq 1}$ son bases ortonormales de $\mathcal{H}$, entonces

$$F_{j,k} = \phi_j \otimes \psi_k$$

es una base ortonormal de S_2.

Demostración. Por las propiedades de la traza, es claro que $\langle S, T \rangle$ es lineal en S y $\langle S, T \rangle = \overline{\langle T, S \rangle}$. Si $T \in S_2$ entonces

$$\langle T, T \rangle^{1/2} = (tr(T^*T))^{1/2} = (tr(|T|^2))^{1/2} = \|T\|_2.$$

Ya que $\|T\|_2$ es una norma bajo la cual S_2 es completo, se sigue que S_2 es un espacio de Hilbert con el producto interno definido.

Dado que

$$F^*_{j,k}F_{l,m} = (\psi_k \otimes \phi_j)(\psi_l \otimes \phi_m) = \langle \phi_l, \phi_j \rangle (\psi_k \otimes \psi_m),$$

tenemos que

$$\langle F_{l,m}, F_{j,k} \rangle = tr(F^*_{j,k}F_{l,m}) = \langle \phi_l, \phi_j \rangle \langle \psi_k, \psi_m \rangle =$$

$$= \begin{cases} 1 & \text{si } l = j \text{ y } m = k \\ 0 & \text{en otro caso .} \end{cases}$$

Entonces $\{F_{j,k} : j, k \in \mathbb{N}\}$ es un sistema ortonormal de S_2. Notar que si $T \in S_2$, entonces

$$\langle T, F_{j,k} \rangle = tr(F^*_{j,k}T) = tr(\psi_k \otimes T^*\phi_j) = \langle T\psi_k, \phi_j \rangle .$$

Si T es ortogonal a cada $F_{j,k}$, entonces $\langle T\psi_k, \phi_j \rangle = 0$, para todo $j, k \in \mathbb{N}$. Dado que T es continuo y una combinación lineal de ψ_k ó ϕ_j es densa en $\mathcal{H}$, se sigue que $\langle Tx, y \rangle = 0$, para cada $x, y \in \mathcal{H}$ y por lo tanto $T = 0$. Entonces $F_{j,k} = \phi_j \otimes \psi_k$ es una base ortonormal de S_2. □

Teorema B.1.7. *Sea $T \in L(\mathcal{H})$ y sean $\{\phi_j : j \in \mathbb{N}\}$, $\{\psi_j : j \in \mathbb{N}\}$ bases ortonormales de $\mathcal{H}$ y $F_{j,k} = \psi_j \otimes \psi_k$. entonces las siguientes condiciones son equivalentes:*

1. $\sum_{j\geq 1} \|T\psi_j\|^2 < \infty$,
2. $\sum_{k\geq 1} \sum_{j\geq 1} |\langle T\psi_j, \phi_k \rangle|^2 < \infty$,
3. $T \in S_2$.

Cuando estas condiciones se satisfacen, se sigue que

$$\|T\|_2^2 = \sum_{j\geq 1} \|T\psi_j\|^2 = \sum_{k\geq 1}\sum_{j\geq 1} |\langle T\psi_j, \phi_k \rangle|^2,$$

y

$$T = \sum_{k\geq 1}\sum_{j\geq 1} \langle T\psi_k, \phi_j \rangle F_{j,k}.$$

Demostración. Ya que

$$\|T\psi_k\|^2 = \sum_{j\geq 1} |\langle T\psi_k, \phi_j \rangle|^2,$$

se sigue que

$$\sum_{k\geq 1} \|T\psi_k\|^2 = \sum_{k\geq 1}\sum_{j\geq 1} |\langle T\psi_k, \phi_j \rangle|^2, \tag{B.3}$$

(ya sea esta suma finita o infinita), por lo tanto 1. y 2. son equivalentes.

Si $T \in S_2$, entonces ya que $\{F_{j,k} : j,k \in \mathbb{N}\}$ es una base ortonormal en S_2 tenemos

$$\|T\|_2^2 = \sum_{j\geq 1,\ k\geq 1} |\langle T, F_{j,k} \rangle|^2 = \sum_{j\geq 1,\ k\geq 1} |\langle T\psi_k, \phi_j \rangle|^2. \tag{B.4}$$

Y además

$$T = \sum_{j\geq 1,\ k\geq 1} \langle T, F_{j,k} \rangle F_{j,k} = \sum_{j\geq 1,\ k\geq 1} \langle T\psi_k, \phi_j \rangle F_{j,k}, \tag{B.5}$$

donde la serie en (B.5), converge con respecto a $\|\cdot\|_2$. En particular, se sigue de (B.4) que 2. se satisface, entonces 3. implica 2.

Recíprocamente, supongamos que $T \in L(\mathcal{H})$ y 2. se satisface. Si $c_{j,k} = \langle T\psi_k, \phi_j \rangle$, tenemos que $\sum_{j\geq 1,\ k\geq 1} |c_{j,k}|^2 < \infty$, entonces

$$\sum_{j\geq 1,\ k\geq 1} c_{j,k} F_{j,k},$$

converge con respecto a $\|\cdot\|_2$ a un elemento $S \in S_2$, para el cual $\langle S, F_{j,k} \rangle = c_{j,k}$. Ya que

$$\langle S\psi_k, \phi_j \rangle = \langle S, F_{j,k} \rangle = c_{j,k} = \langle T\psi_k, \phi_j \rangle,$$

para cada $j,k \in \mathbb{N}$, se sigue que $T = S \in S_2$, y entonces 2. implica 3.

Se ha mostrado, que las condiciones 1., 2. y 3. son equivalentes, y que dada dichas condiciones, las ecuaciones (B.3), (B.4) y (B.5) se satisfacen. $\square$

BIBLIOGRAFÍA

[1] J. G Aiken, J. A Erdos, and J. A Goldstein. On Löwdin orthogonalization. *International Journal of Quantum Chemistry*, 18(4):1101–1108, 1980.

[2] J. G. Aiken, J. A. Erdos, and J. A. Goldstein. Unitary approximation of positive operators. *Illinois Journal of Mathematics*, 24(1):61–72, 1980.

[3] W. N. Anderson and G. E. Trapp. Shorted operators. *SIAM J. Appl. Math*, (20):520–525, 1971.

[4] W. N. Anderson Jr. and G. E. Trapp. Shorted operators II. *SIAM J. Appl. Math.*, 28:60–71, 1975.

[5] T. Ando. Generalized Schur complements. *Linear Algebra Appl.*, 27:173–186, 1979.

[6] J. Antezana, G. Corach, M. Ruiz, and D. Stojanoff. Oblique projections and frames. *Proceedings of the American Mathematical Society*, 134(4):1031–1037, 2006.

[7] M. L. Arias and M. C. Gonzalez. Positive solutions to operator equations $AXB = C$. *Linear Algebra and its Applications*, 433(6):1194–1202, 2010.

[8] A. Ben-Israel and T. N. E. Greville. *Generalized inverse: Theory and Applications*. Springer-Verlag, 2003.

[9] R. Bouldin. The product of operators with closed range. *Tohoku Math. J.*, 25:359–363, 1973.

[10] D. Carlson. What are Schur complements, anyway? *Linear Algebra Appl.*, (74):257–275, 1986.

[11] Y. Censor and T. Elfving. Block-iterative algorithms with diagonally scaled oblique projections for the linear feasibility problem. *SIAM Journal on Matrix Analysis and Applications*, 24(1):40–58, 2002.

[12] Y. Censor and T. Elfving. Iterative algorithms with seminorm-induced oblique projections. In *Abstract and Applied Analysis*, volume 2003, pages 387–406, 2003.

[13] Y. Censor, D. Gordon, and R. Gordon. Bicav: A block-iterative parallel algorithm for sparse systems with pixel-related weighting. *Medical Imaging, IEEE Transactions on*, 20(10):1050–1060, 2001.

[14] Y. Censor, D. Gordon, and R. Gordon. Component averaging: An efficient iterative parallel algorithm for large and sparse unstructured problems. *Parallel computing*, 27(6):777–808, 2001.

[15] Y. Changsen. On the critical points of the map $F_p : X \to \|AXB - C\|_p^p$. *Applied Mathematics and Mechanics*, 21(4):485–488, 2000.

[16] J. S. Chipman. On least squares with insufficient observations. *J. Amer. Statist. Assoc.*, 59:1078–1111, 1964.

[17] M. Contino, J. I. Giribet, and A. Maestripieri. Weighted least square solutions of the equation $AXB - C = 0$. *Linear Algebra and its Applications*, 518:177–197, 2017.

[18] M. Contino, J. I. Giribet, and A. Maestripieri. Weighted Procrustes problems. *Journal of Mathematical Analysis and Applications*, 445(1):443–458, 2017.

[19] J. B. Conway. *A course in functional analysis*. Springer Science & Business Media, 2013.

[20] G. Corach, G. Fongi, and A. Maestripieri. Weighted projections into closed subspaces. *preprint arXiv:1305.6357*, 2013.

[21] G. Corach, M. C. Gonzalez, and A. Maestripieri. Unbounded symmetrizable idempotents. *Linear Algebra and its Applications*, 437(2):659–674, 2012.

[22] G. Corach and A. Maestripieri. Weighted generalized inverses, oblique projections and least squares problems. *Numer. Funct. Anal. Optim.*, 26(6):659–673, 2005.

[23] G. Corach, A. Maestripieri, and D. Stojanoff. Generalized orthogonal projections and shorted operators. *Margarita mathematica, Univ. La Rioja, España*, pages 607–625, 2001.

[24] G. Corach, A. Maestripieri, and D. Stojanoff. Oblique projections and Schur complements. *Acta Sci. Math. (Szeged)*, 67:337–256, 2001.

[25] G. Corach, A. Maestripieri, and D. Stojanoff. Generalized Schur complements and oblique projections. *Linear Algebra Appl.*, 341:259–272, 2002.

[26] G. Corach, A. Maestripieri, and D. Stojanoff. Oblique projections and abstract splines. *J. Approx. Theory*, pages 189–206, 2002.

[27] G. Corach, A. Maestripieri, and D. Stojanoff. A classification of projectors. *Banach Center Publ., Institute of Mathematics, Polish Academy of Sciences, Warszawa,*, 67:145–160, 2004.

[28] G. Corach, A. Maestripieri, and D. Stojanoff. Projections in operator ranges. *Proc. Amer. Math. Soc.*, 134(3):765–778, 2006.

[29] G. Corach, P. Massey, and M. Ruiz. Procrustes problems and Parseval quasi-dual frames. *Acta Applicandae Mathematicae*, 131(1):179–195, 2014.

[30] R. W. Cottle. Manifestations of the Schur complements. *Linear Algebra Appl.*, 8:189–211, 1974.

[31] Y. Dai, H. Li, and M. He. Element-wise factorization for n-view projective reconstruction. In *Computer Vision–ECCV 2010*, pages 396–409. Springer, 2010.

[32] F. Deutsch. The angle between subspaces in Hilbert spaces. *Aproximation theory, wavelets and applications*, pages 107–130, 1995.

[33] J. Diestel. Geometry of Banach spaces. *Lecture Notes in Mathematics*, 485, 1975.

[34] J. Dixmier. Étude sur less variétés et les opérateurs de julia avec quelques applications. *Bull. Soc. Math. France*, 77:11–101.

[35] J. Dixmier. Position relative de deux variétés linéaires fermées dans un espace de Hilbert. *Revue Scientifique*, 86:387–399, 1948.

[36] S. M. Djouadi. On the optimality of the proper orthogonal decomposition and balanced truncation. In *Decision and Control, 2008. CDC 2008. 47th IEEE Conference on*, pages 4221–4226. IEEE, 2008.

[37] R. G. Douglas. On majorization, factorization and range inclusion of operators in Hilbert space. 17:413–416, 1966.

[38] Y. C. Eldar and T. Werther. General framework for consistent sampling in Hilbert spaces. *International Journal of Wavelets, Multiresolution and Information Processing*, 3(04):497–509, 2005.

[39] L. Eldén. Perturbation theory for the least squares problem with linear equality contraints. *SIAM J. Numer. Anal.*, 17(3):338–350, 1980.

[40] H. W Engl and M. Z. Nashed. New extremal characterizations of generalized inverses of linear operators. *Journal of Mathematical Analysis and Applications*, 82(2):566–586, 1981.

[41] J.A. Erdös. On the trace of a trace class operator. *Bulletin of the London Mathematical Society*, 6(1):47–50, 1974.

[42] A. Eriksson and A. Van Den Hengel. Efficient computation of robust low-rank matrix approximations in the presence of missing data using the l 1 norm. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 771–778. IEEE, 2010.

[43] M. Frank, V. Paulsen, and T. Tiballi. Symmetric approximation of frames and bases in Hilbert spaces. *Transactions of the American Mathematical Society*, 354(2):777–793, 2002.

[44] S. H. Friedberg, A. J. Insel, and L. E. Spence. *Algebra lineal*. Publicaciones Cultural, 1982.

[45] K. Friedrichs. On certain inequalities and characteristic value problems for analytic fuctions and for functions of two variables. *Trans. Amer. Math. Soc.*, 41:321–364, 1937.

[46] I. Gohberg and M. G. Kreĭn. *Introduction to the theory of linear nonselfadjoint operators*, volume 18. American Mathematical Soc., 1969.

[47] G. Goldstein and J. A. Goldstein. The best generalized inverse. *Journal of mathematical analysis and applications*, 252(1):91–101, 2000.

[48] J. A. Goldstein and M. Levy. Hilbert-Schmidt approximation problems arising in quantum chemistry. *Advances in Applied Mathematics*, 5(2):216–225, 1984.

[49] J. A Goldstein and M. Levy. Linear algebra and quantum chemistry. *American Mathematical Monthly*, 98(10):710–718, 1991.

[50] T. N. E Greville. Note on fitting of functions of several independent variables. *Journal of the Society for Industrial and Applied Mathematics*, 9(1):109–115, 1961.

[51] D. Han. Frame representations and Parseval duals with applications to gabor frames. *Transactions of the American Mathematical Society*, 360(6):3307–3326, 2008.

[52] E. Haynsworth. Determination of the inertia of a partitioned hermitian matrix. *Linear Algebra Appl.*, 1:73–81, 1968.

[53] S. Izumino. The product of operadors with closed range and an extension of the reverse order law. *Tohoku Math. J*, 34:43–52, 1982.

[54] D. J. Keckic. Orthogonality in S_1 and S_∞ spaces and normal derivations. *J. Operator Theory*, 51:89–104, 2004.

[55] M. G. Krein. The theory of self-adjoint extensions of semibounded hermitian operators and its applications. *Mat. Sb. (N. S.)*, 20(62):431–495, 1947.

[56] I. N. Levine. *Química cuántica*. Pearson Educación, 2001.

[57] I. N. Levine. *Quantum chemistry*, volume 6. Pearson Prentice Hall Upper Saddle River, NJ, 2009.

[58] P. O. Löwdin. On the nonorthogonality problem. *Advances in quantum chemistry*, 5:185, 1970.

[59] P. J. Maher. The nearest self-adjoint operator. *The Journal of chemical physics*, 92(11):6978–6978, 1990.

[60] P. J. Maher. Some norm inequalities concerning generalized inverses. *Linear algebra and its applications*, 174:99–110, 1992.

[61] P. J. Maher. Some matrix approximation problems arising from quantum chemistry. *Proc. Indian Nat. Sci. Acad., Part A*, 64:715–724, 1998.

[62] P.J. Maher. Some operator inequalities concerning generalized inverses. *Illinois Journal of Mathematics*, 34(3):503–514, 1990.

[63] P.J. Maher. Some norm inequalities concerning generalized inverses, 2. *Linear algebra and its applications*, 420(2):517–525, 2007.

[64] C. A. MCCarthy. C_p. *Israel J. Math.*, 5:249–271, 1967.

[65] L. Mirsky. Symmetric gauge functions and unitarily invariant norms. *The quarterly journal of mathematics*, 11(1):50–59, 1960.

[66] S. K. Mitra and C. R. Rao. Projections under seminorms and generalized Moore Penrose inverses. *Linear Algebra and Its Applications*, 9:155–167, 1974.

[67] M. Z. Nashed. Inner, outer, and generalized inverses in Banach and Hilbert spaces. *Numer. Funct. Anal. Optim.*, 9(3-4):261–325, 1987.

[68] M. Z. Nashed and G. F. Votruba. *A unified operator theory of generalized inverses*, volume 32. Academic Press, New York, 1976.

[69] Z. Pasternak-Winiarski. On the dependence of the orthogonal projector on deformations of the scalar product. *Studia Mathematica*, 128:1, 1998.

[70] G. K. Pedersen and R. V. Kadison. Means and convex combinations of unitary operators. *Math. Scand.*, 57:249–266, 1985.

[71] E. L. Pekarev. Shorts of operators and some extremal problems. *Acta Sci. Math. (Szeged)*, 56:147–163, 1992.

[72] C. R. Rao and S. K. Mitra. *Generalized inverse of a matrix and its applications*. Wiley and Sons, New York, 1971.

[73] C. R Rao and S. K. Mitra. Theory and application of constrained inverse of matrices. *SIAM Journal on Applied Mathematics*, 24(4):473–488, 1973.

[74] F. Riesz and B. Szőkefalvi-Nagy. *Functional analysis*. Ungar, 1960.

[75] J. R. Ringrose. *Compact non-self-adjoint operators*. Van Nostrand-Reinhold, New York, 1971.

[76] H. L. Royden and P. Fitzpatrick. *Real analysis*, volume 198. Macmillan New York, 1988.

[77] W. Rudin. *Functional analysis 2nd. ed.* New York, McGraw-Hill, 1991.

[78] R. Schatten. *Norm ideals of completely continuous operators*. Springer, 1960.

[79] R. Schatten and J. von Neumann. The cross-space of linear transformations. II. *Annals of Mathematics*, pages 608–630, 1946.

[80] J. R. Singler. Optimality of balanced proper orthogonal decomposition for data reconstruction. *Numerical functional analysis and optimization*, 31(7):852–869, 2010.

[81] T. R. Tiballi. *Symmetric orthogonalization of vectors in Hilbert spaces, Ph.D. Thesis*. University of Houston, Houston Texas, U.S.A., 1991.

[82] M. Unser. Sampling-50 years after Shannon. *Proceedings of the IEEE*, 88(4):569–587, 2000.

[83] M. Unser and A. Aldroubi. A general sampling theory for nonideal acquisition devices. *IEEE Trans. on Signal Proc.*, 42(11):2915–2925, 1994.

[84] J. Van Tiel. *Convex analysis: An introduction text*. Wiley, New York, 1984.

[85] G. S. Watson. Linear least squares regression. *The Annals of Mathematical Statistics*, pages 1679–1699, 1967.

[86] G. Zyskind. On canonical forms, non-negative covariance matrices and best and simple least squares linear estimators in linear models. *The Annals of Mathematical Statistics*, pages 1092–1109, 1967.