

El algoritmo alternante de von Neumann

Ana Crespo Gallego



Un trabajo fin de grado presentado con
la supervisión de Miguel Lacruz Martín

Departamento de Análisis Matemático
Facultad de Matemáticas
Universidad de Sevilla
10 de septiembre de 2020

Índice

Resumen	3
1 Introducción	5
2 El método para dos subespacios	7
2.1 Convergencia general	7
2.2 Ritmo de convergencia	10
3 El método para varios subespacios	21
3.1 El algoritmo de Dykstra	21
3.2 Convergencia geométrica	29
4 Diversas aplicaciones del método	33
4.1 Sistemas de ecuaciones lineales	33
4.2 Sistema de inecuaciones lineales	34
4.3 Restauración de imágenes	34
4.4 Solución al problema de Dirichlet	35
4.5 Núcleos de Bergman	36
4.6 Otras aplicaciones	37
5 Apéndice	39
5.1 Resultados auxiliares conocidos del grado	39
5.2 Teoría elemental de las álgebras de Banach	41
5.3 El teorema espectral para operadores hermíticos	45

Resumen

The aim of this work is to study in depth the alternating algorithm proposed by John von Neumann in 1933. This algorithm deals with the problem of finding the orthogonal projection of a point in a Hilbert space onto the intersection of two closed subspaces, in terms of the alternating orthogonal projections onto the respective subspaces.

We will study the convergence of the method and its rate of convergence, as well as the generalization of such algorithm to more than two subspaces or to the case when the subspaces are replaced by convex subsets. Finally, we discuss some applications where the method plays an important role, like the approximate solution of systems of linear equations or the Dirichlet problem.

El objetivo de este trabajo es estudiar en profundidad el algoritmo alternante propuesto por John von Neumann en 1933. Este algoritmo trata el problema de encontrar la proyección ortogonal de un punto en un espacio de Hilbert sobre la intersección de dos subespacios cerrados, en términos de la proyecciones ortogonales alternadas sobre los respectivos subespacios. Además estudiamos la convergencia del método y su ritmo de convergencia, así como la generalización de dicho algoritmo a más de dos subespacios, o el caso cuando los subespacios se reemplazan por subconjuntos convexos. Por último, discutimos algunas aplicaciones donde el método desempeña un papel importante, como la aproximación de sistemas de ecuaciones lineales o el problema de Dirichlet.

Capítulo 1

Introducción

El método alternante fue propuesto en 1933 por John von Neumann, que trató el problema de encontrar la proyección de un punto dado en un espacio de Hilbert sobre la intersección de dos subespacios cerrados.

Comenzaremos con el capítulo 2 describiendo dicho resultado obtenido por von Neumann [3], que da nombre al trabajo, y veremos qué ocurre con la convergencia de dicho método iterativo en su formulación más simple (intersección de dos subespacios). Observaremos que este ritmo depende del ángulo entre los subespacios involucrados; cuando el ángulo se encuentra entre $[0, 1)$ la convergencia es geométrica y, si en el caso contrario, el ángulo es 1, habla de convergencia arbitrariamente lenta.

El capítulo 3 consiste en probar el algoritmo de Dykstra [3], que estudia las mejores aproximaciones de la intersección de un número finito de conjuntos convexos cerrados. Además de extender el teorema de von Neumann a un número finito de subespacios cerrados. Dicho resultado fue probado por Halperin. Además veremos que la convergencia para una familia finita de subespacios cerrados con ángulo $\alpha \in [0, 1)$ sigue siendo geométrica.

En el capítulo 4 conoceremos diferentes áreas de las matemáticas donde el método de von Neumann ha desempeñado un papel importante, en las que se incluyen la resolución de ecuaciones e inecuaciones lineales [3], solución al problema de Dirichlet [2] y restauración de imágenes [4], entre otras.

Por último, en el capítulo 5 daremos definiciones y resultados que son importantes para la preparación del trabajo, algunos conocidos en el grado y otros no, como el teorema espectral [5] y la teoría elemental de las álgebras de Banach [7].

Capítulo 2

El método para dos subespacios

En este capítulo vamos a estudiar la formulación del método más simple, es decir, la intersección de dos subespacios. Veremos este caso primero para motivar los resultados más generales que vendrán después. Este capítulo está basado en el libro [3].

2.1. Convergencia general

Notación 2.1.1. $P_i = P_{M_i}$ para $i = 1, 2$.

Teorema 2.1.2 (Teorema de von Neumann). *Sea H un espacio de Hilbert, M_1 y M_2 dos subespacios cerrados de H y P_1 y P_2 las proyecciones ortogonales sobre M_1 y M_2 , respectivamente. Entonces, para cada $x \in H$,*

$$\lim_{n \rightarrow \infty} (P_2 P_1)^n(x) = P_{M_1 \cap M_2}(x)$$

Demostración. Sea $x \in H$ y (x_n) una sucesión de vectores definida como:

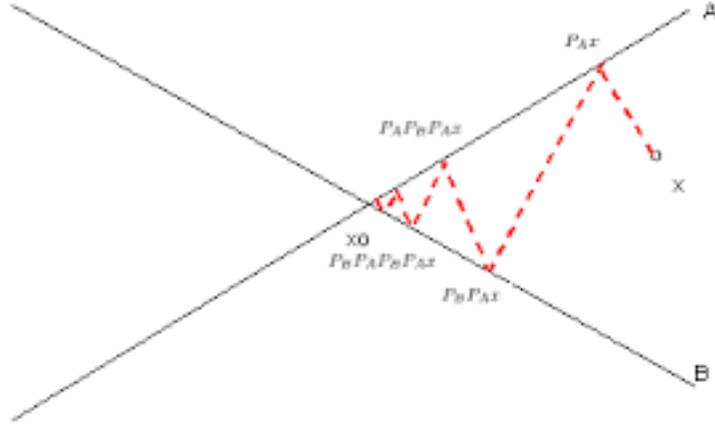
$$\begin{aligned} x_0 &= x \\ x_{2n-1} &= P_1(x_{2n-2}) = P_1(P_2 P_1)^{(n-1)}x \quad \forall n = 1, 2, \dots \\ x_{2n} &= P_2(x_{2n-1}) = (P_2 P_1)^n x. \end{aligned}$$

Tenemos que probar que $x_{2n} \rightarrow P_{M_1 \cap M_2}(x)$, por lo que (x_n) convergería a $P_{M_1 \cap M_2}(x)$, ya que $x_{2n+1} = P_1 x_{2n}$ y también convergería a $P_{M_1 \cap M_2}(x)$.

Como $\|P_i\| \leq 1$, entonces

$$\|x_{2n}\| = \|P_2 x_{2n-1}\| \leq \|x_{2n-1}\| = \|P_1 x_{2n-2}\| \leq \|x_{2n-2}\|.$$

Figura 2.1: Representación del algoritmo de von Neumann



La sucesión de la norma ($\|x_n\|$) es decreciente y por tanto, existe λ tal que

$$\lambda = \lim_{n \rightarrow \infty} \|x_n\|.$$

Como los P_i son autoadjuntos, $x_{2n} \in M_2$ y $x_{2n-1} \in M_1$ y se deduce que

$$\begin{aligned} \langle x_{2n}, x_{2m} \rangle &= \langle P_2 x_{2n-1}, x_{2m} \rangle \\ &= \langle x_{2n-1}, P_2 x_{2m} \rangle \\ &= \langle x_{2n-1}, x_{2m} \rangle \\ &= \langle P_1 x_{2n-1}, x_{2m} \rangle \\ &= \langle x_{2n-1}, P_1 x_{2m} \rangle \\ &= \langle x_{2n-1}, x_{2m+1} \rangle \end{aligned}$$

$$\begin{aligned} \langle x_{2n+1}, x_{2m+1} \rangle &= \langle P_1 x_{2n}, x_{2m+1} \rangle \\ &= \langle x_{2n}, P_1 x_{2m+1} \rangle \\ &= \langle x_{2n}, x_{2m+1} \rangle \\ &= \langle P_2 x_{2n}, x_{2m+1} \rangle \\ &= \langle x_{2n}, P_2 x_{2m+1} \rangle \\ &= \langle x_{2n}, x_{2m+2} \rangle. \end{aligned}$$

Por tanto, tenemos que

$$\langle x_{2n}, x_{2m} \rangle = \langle x_{2n-1}, x_{2m+1} \rangle$$

$$\langle x_{2n+1}, x_{2m+1} \rangle = \langle x_{2n}, x_{2m+2} \rangle.$$

Combinando ambos resultados anteriores, podemos deducir que si $n \geq m$, entonces

$$\langle x_{2n}, x_{2m} \rangle = \langle x_{2n-1}, x_{2m+1} \rangle = \langle x_{2n-2}, x_{2m+2} \rangle = \cdots = \langle x_{2n-k}, x_{2m+k} \rangle$$

para algún $0 \leq k \leq 2n$.

Para $k = n - m$, se obtiene

$$\langle x_{2n}, x_{2m} \rangle = \langle x_{n+m}, x_{n+m} \rangle = \|x_{n+m}\|^2.$$

Supongamos ahora que $m \leq n$

$$\|x_{2n} - x_{2m}\|^2 = \|x_{2n}\|^2 - 2\langle x_{2n}, x_{2m} \rangle + \|x_{2m}\|^2 = \|x_{2n}\|^2 - 2\|x_{n+m}\|^2 + \|x_{2m}\|^2.$$

Cuando $m \rightarrow \infty$, tenemos $\lambda^2 - 2\lambda^2 + \lambda^2 = 0$.

Esto demuestra que (x_{2n}) es una sucesión de Cauchy en M_2 . Como H es completo y M_2 es cerrado, entonces existe $y \in M_2$ tal que $x_{2n} \rightarrow y$.

Además:

$$\langle x_{2n}, x_{2n-1} \rangle = \langle P_2 x_{2n}, x_{2n-1} \rangle = \langle x_{2n}, P_2 x_{2n-1} \rangle = \langle x_{2n}, x_{2n} \rangle = \|x_{2n}\|^2.$$

Esto implica que

$$\begin{aligned} \|x_{2n} - x_{2n-1}\|^2 &= \|x_{2n}\|^2 - 2\langle x_{2n}, x_{2n-1} \rangle + \|x_{2n-1}\|^2 \\ &= \|x_{2n}\|^2 - 2\|x_{2n}\|^2 + \|x_{2n-1}\|^2 = \|x_{2n-1}\|^2 - \|x_{2n}\|^2. \end{aligned}$$

Cuando $n \rightarrow \infty$, tenemos $\lambda^2 - \lambda^2 = 0$ y se deduce que

$$\|x_{2n-1} - y\| \leq \|x_{2n-1} - x_{2n}\| + \|x_{2n} - y\| \rightarrow 0.$$

Ya que $x_{2n-1} \in M_1 \forall n$ y M_1 es cerrado, entonces $y \in M_1$.

Hemos probado que

$$\lim_{n \rightarrow \infty} x_n = y \in M_1 \cap M_2.$$

Para ver que $y = P_{M_1 \cap M_2}(x)$ es suficiente probar que $x - y \perp M_1 \cap M_2$.

Sea $z \in M_1 \cap M_2$, entonces:

$$\begin{aligned}\langle x_{2n}, z \rangle &= \langle x_{2n} - x_{2n-1} + x_{2n-1}, z \rangle \\ &= \langle P_1 x_{2n-1} - x_{2n-1} + x_{2n-1}, z \rangle \\ &= \langle x_{2n-1}, z \rangle.\end{aligned}$$

De forma similar tenemos que

$$\langle x_{2n-1}, z \rangle = \langle x_{2n-2}, z \rangle.$$

Por inducción

$$\langle x_{2n}, z \rangle = \langle x_0, z \rangle = \langle x, z \rangle.$$

Entonces

$$\langle x - y, z \rangle = \langle x, z \rangle - \langle y, z \rangle = \langle x, z \rangle - \lim_{n \rightarrow \infty} \langle x_{2n}, z \rangle = \langle x, z \rangle - \langle x, z \rangle = 0.$$

□

2.2. Ritmo de convergencia

El ritmo de convergencia del método de proyecciones alternas es rápida si el ángulo entre los subespacios se encuentra entre $[0, 1)$ o arbitrariamente lenta si dicho ángulo es 1.

Definición 2.2.1 (Ángulo entre dos subespacios). El ángulo entre dos subespacios M_1 y M_2 es el ángulo $\alpha(M_1, M_2) \in [0, \pi/2]$ cuyo coseno $c(M_1, M_2) = \cos \alpha(M_1, M_2)$ es definido por la expresión

$$c(M_1, M_2) := \sup\{|\langle x, y \rangle| : x \in M_1 \cap (M_1 \cap M_2)^\perp, \|x\| \leq 1, y \in M_2 \cap (M_1 \cap M_2)^\perp, \|y\| \leq 1\}.$$

2.2.1. Convergencia geométrica

Teorema 2.2.2. Sean M_1 y M_2 dos subespacios cerrados en un espacio de Hilbert H y $c = c(M_1, M_2)$. Entonces para cada $x \in H$,

$$\|(P_{M_2} P_{M_1})^n(x) - P_{M_1 \cap M_2}(x)\| \leq c^{2n-1} \|x\| \quad (2.1)$$

para $n = 1, 2, \dots$. Además la constante c^{2n-1} es la más pequeña posible independientemente de x . (Esto lo veremos más adelante en el teorema 3.2.3)

Demostración. Sea $M = M_1 \cap M_2$, $P_i = P_{M_i}$ para $i = 1, 2$ Y $P = P_2 P_1$.
Vamos a probar primero que

$$P_i(M)^\perp \subset M^\perp \quad (2.2)$$

para $i = 1, 2$ y por lo tanto

$$P(M)^\perp \subset M^\perp. \quad (2.3)$$

Probemos (2.2). Sea $x \in M^\perp$ e $y \in M$. Como P_i es autoadjunto e idempotente,

$$\langle P_i(x), y \rangle = \langle x, P_i(y) \rangle = \langle x, y \rangle = 0.$$

Así $P_i(x) \in M^\perp$, que demuestra lo que queríamos probar.

Continuamos viendo que

$$P(M^\perp) = P_2 P_1(M^\perp) \subset P_2(M^\perp) \subset M^\perp.$$

Por lo que hemos probado (2.3). A continuación, tenemos que demostrar

$$\|P(x)\| \leq c \|P_1(x)\| \leq c \|x\|. \quad (2.4)$$

Para ver esto, sea $x \in M^\perp$. Entonces

$$\begin{aligned} \|P(x)\|^2 &= \langle P(x), P(x) \rangle \\ &= \langle P(x) - P_1(x) + P_1(x), P(x) \rangle \\ &= \langle P_1(x), P(x) \rangle \quad (\text{ya que } P(x) - P_1(x) \in M_2^\perp, P(x) \in M_2.) \\ &\leq c \|P_1(x)\| \|P(x)\| \\ &\leq c \|x\| \|P(x)\| \quad (\text{ya que } \|P_i\| \leq 1). \end{aligned}$$

Dividiendo por $\|P(x)\|$ nos da el resultado (2.4).

Ahora supongamos $x \in P(M^\perp)$, entonces $x \in M_2 \cap M^\perp$ por (2.3). Luego $P_1(x) \in M_1 \cap M^\perp$ por (2.2) y obtenemos

$$\|P_1(x)\|^2 = \langle P_1(x), P_1(x) \rangle = \langle x, P_1(x) \rangle = c \|x\| \|P_1(x)\|.$$

Por tanto $\|P_1(x)\| \leq c \|x\|$. Sustituyendo esto en (2.4), obtenemos

$$\|P(x)\| \leq c^2 \|x\| \quad \forall x \in P(M^\perp). \quad (2.5)$$

A partir de (2.3), (2.4) y una inducción en (2.5), tenemos que

$$\begin{aligned} \|(P_2 P_1)^n(x) - P_M(x)\| &= \|P^n(x) - P^n P_M(x)\| \\ &= \|P^n[x - P_M(x)]\| \\ &= \|P^{n-1}[P(x - P_M(x))]\| \\ &\leq c^{2n-1} \|P(x - P_M(x))\| \\ &\leq c^{2n-1} \|x - P_M(x)\| \\ &\leq c^{2n-1} \|x\|. \end{aligned}$$

Por lo que hemos probado (2.1), lo que verifica el teorema. □

2.2.2. Convergencia arbitrariamente lenta

En esta parte del capítulo hablaremos lo que ocurriría si la convergencia no fuese geométrica, lo que sucedería si la suma de ambos subespacios no fuese cerrada, es decir, si el ángulo entre dichos subespacios fuese 1. Esta sección está basada en el artículo [1].

Lema 2.2.3. *Son equivalentes:*

- (1) $c(M_1, M_2) < 1$.
- (2) $M_1 \cap (M_1 \cap M_2)^\perp + M_2 \cap (M_1 \cap M_2)^\perp$ es cerrado.
- (3) $M_1^\perp + M_2^\perp$ es cerrado.
- (4) $M_1 + M_2$ es cerrado.

Demostración. Sea $Y := M_1 \cap (M_1 \cap M_2)^\perp$ y $Z := M_2 \cap (M_1 \cap M_2)^\perp$. Primero probemos el siguiente lema:

Lema 2.2.4.

- (a) $Y + Z = (M_1 + M_2) \cap (M_1 \cap M_2)^\perp$.
- (b) $M_1 + M_2 = Y + Z + M_1 \cap M_2$.

Demostración.

(a). Notemos que $Y + Z \subset (M_1 + M_2) \cap (M_1 \cap M_2)^\perp$. Sea $x \in (M_1 + M_2) \cap (M_1 \cap M_2)^\perp$, luego $x = y + z$, donde $y \in M_1$ y $z \in M_2$. Ya que $x \in (M_1 \cap M_2)^\perp$, se deduce que $P_{M_1 \cap M_2} x = 0$. Por tanto

$$\begin{aligned} x &= x - P_{M_1 \cap M_2} x = (y - P_{M_1 \cap M_2} y) + (z - P_{M_1 \cap M_2} z) \\ &\in M_1 \cap (M_1 \cap M_2)^\perp + M_2 \cap (M_1 \cap M_2)^\perp = Y + Z \end{aligned}$$

usando la caracterización de mejor aproximación de subespacios. Lo que prueba (a).

(b). Observamos que $Y + Z + M_1 \cap M_2 \subset M_1 + M_2$. Ahora fijemos un $x \in M_1 + M_2$, luego $x = y + z$ para algún $y \in M_1$ y $z \in M_2$.

$$x = P_{M_1 \cap M_2}(x) + P_{(M_1 \cap M_2)^\perp}(x) = P_{M_1 \cap M_2}(y) + P_{(M_1 \cap M_2)^\perp}(y) + P_{M_1 \cap M_2}(z) + P_{(M_1 \cap M_2)^\perp}(z).$$

Pero como $M_1 \cap M_2 \subset M_1$, $P_{M_1 \cap M_2}(x)$ conmuta con P_M , se deduce que $P_{(M_1 \cap M_2)^\perp}(x) = I - P_{M_1 \cap M_2}(x)$ también conmuta con P_{M_1} . Además $P_{(M_1 \cap M_2)^\perp}(y) \in M_1 \cap (M_1 \cap M_2)^\perp = Y$ y $P_{(M_1 \cap M_2)^\perp}(z) \in M_2 \cap (M_1 \cap M_2)^\perp = Z$. Esto prueba que $x \in M_1 \cap M_2 + Y + Z$ y verifica (b). □

(2) \Leftrightarrow (3). Si $M_1 + M_2$ es cerrado, $(M_1 \cap M_2)^\perp$ es también cerrado. Por lo que $Y + Z$ es cerrado por lema 2.2.4 (a).

Para probarlo en el otro sentido, supongamos $Y + Z$ cerrado. Ya que $Y + Z \subset (M_1 \cap M_2)^\perp$, la suma $(Y + Z) + M_1 \cap M_2$ es cerrada porque es la suma de subespacios ortogonales cerrados. Usando el lema 2.2.4 (b), deducimos que $M_1 + M_2$ cerrado.

(1) \Rightarrow (2) Sea $c = c(M_1, M_2) < 1$. Tenemos que probar que $Y + Z$ es cerrado. Además

$$|\langle y, z \rangle| \leq c \|y\| \|z\| \quad \forall y \in Y \quad \forall z \in Z.$$

Ahora sea $x_n \in Y + Z$ y $x_n \rightarrow x$. Es suficiente probar que $x \in Y + Z$. Sea $x_n = y_n + z_n$ con $y_n \in Y$ y $z_n \in Z$. Luego $\{x_n\}$ es acotado, así que existe una constante $\rho > 0$ tal que

$$\begin{aligned} \rho &\geq \|x_n\|^2 = \|y_n + z_n\|^2 \\ &= \|y_n\|^2 + 2\langle y_n, z_n \rangle + \|z_n\|^2 \\ &\geq \|y_n\|^2 - 2\langle y_n, z_n \rangle + \|z_n\|^2 \\ &\geq \|y_n\|^2 - 2c\|y_n\| \|z_n\| + \|z_n\|^2 \\ &= (\|y_n\| - \|z_n\|)^2 + 2(1 - c)\|y_n\| \|z_n\|. \end{aligned}$$

Ya que $0 \leq c < 1$, las sucesiones $\{\|y_n\| - \|z_n\|\}$ y $\{\|y_n\| \|z_n\|\}$ son acotadas. Por tanto, $\{\|y_n\|\}$ y $\{\|z_n\|\}$ son acotadas. Al pasar a una subsucesión, $\{y_n\}$ converge débilmente a algún y , y $\{z_n\}$ converge débilmente a algún z . Por tanto, $y \in Y$, $z \in Z$ y $x_n = y_n + z_n \rightarrow y + z$ débilmente. Así que $x_n \rightarrow x$ implica que $x_n \rightarrow x$ débilmente, y por lo tanto, $x = y + z \in Y + Z$. Esto prueba (2).

(2) \Rightarrow (1). Suponemos que $Y + Z$ es cerrado. Luego $H := Y + Z$ es un espacio de Hilbert y $Y \cap Z = \{0\}$. La aplicación lineal $Q : H \rightarrow Y$ definida por

$$Q(y + z) = y, \quad y \in Y, \quad z \in Z,$$

es continua.

Si (1) no fuese cierto, es decir, si $c := c(M_1, M_2) = 1$ y existiera $y_n \in Y$, $z_n \in Z$ con $\|y_n\| = 1 = \|z_n\|$ y $1 = c = \lim_{n \rightarrow \infty} \langle x, y \rangle$. Luego,

$$\|y_n - z_n\|^2 = \|y_n\|^2 - 2\langle y_n, z_n \rangle + \|z_n\|^2 = 2 - 2\langle y_n, z_n \rangle \rightarrow 0$$

implica que

$$y_n = Q(y_n - z_n) \rightarrow Q(0) = 0,$$

lo cual es absurdo. Por lo tanto, $c < 1$ y esto prueba (1).

Veamos ahora que $M_1 + M_2$ es cerrado si y solo si $M_1^\perp + M_2^\perp$ también lo es. Como $(M_1 \cap M_2)^\perp = M_1^\perp + M_2^\perp$, si $(M_1 \cap M_2)^\perp$ es cerrado, $M_1^\perp + M_2^\perp$ es cerrado siempre que $M_1 + M_2$ sea cerrado.

Veamos la otra implicación. $M_1^\perp + M_2^\perp$ es cerrado, por el mismo razonamiento anterior $M_1^{\perp\perp} + M_2^{\perp\perp}$ es cerrado y como $M_1^{\perp\perp} = M_1$ y $M_2^{\perp\perp} = M_2$, entonces $M_1 + M_2$ es cerrado también. □

En este método, para cada punto de partida hay convergencia geométrica o arbitrariamente lenta.

Teorema 2.2.5. (*Dicotomía*) Sea H un espacio de Hilbert, M_1 y M_2 dos subespacios cerrados y $M = M_1 \cap M_2$. Entonces se cumple exactamente una de las siguientes alternativas:

(a). $M_1^\perp + M_2^\perp$ es cerrado. Entonces para cada $x \in H$, $(P_{M_2}P_{M_1})^n(x)$ converge geométricamente a $P_M(x)$ con velocidad $[c(M_1, M_2)]^2$.

(b). $M_1^\perp + M_2^\perp$ no es cerrado. Entonces para cada $x \in H$, $(P_{M_2}P_{M_1})^n(x)$ converge a $P_M(x)$ pero dicha convergencia es arbitrariamente lenta, es decir, para cada sucesión (λ_n) de números reales positivos con $1 > \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \rightarrow 0$, entonces existe un punto $x_\lambda \in H$ tal que $\|(P_{M_2}P_{M_1})^n(x_\lambda) - P_M(x_\lambda)\| \geq \lambda_n$ para todo n .

Demostración. La prueba de (a) es consecuencia inmediata del teorema 2.2.2 y del lema 2.2.3. Ahora probemos (b). Supongamos que $M_1 + M_2$ no es cerrado y que (λ_n) es una sucesión con $1 > \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n > 0$ y $\lambda_n \rightarrow 0$.

Sea $A = M_1 \cap M^\perp$ y $B = M_2 \cap M^\perp$. A y B son dos subespacios cerrados con $A \cap B = \{0\}$. Claramente, $c(A, B) = c(M_1, M_2) = 1$, y por lo tanto $A + B$ no es cerrado (por el lema 2.2.4). Dado que $c(A, B) = \|P_B P_A\|$, se deduce que $\|P_B P_A\| = 1$.

Lema 2.2.6. $T = P_A P_B P_A$ es un operador lineal, acotado y positivo en H y $\|T\| = 1$. Entonces existe un espacio de medida finita (Ω, μ) , una función medible no negativa esencialmente acotada F en Ω y un operador unitario $U : H \rightarrow L_2 := L_2(\Omega, \mu)$ tal que $UTU^{-1} = D$ y $\|F\|_\infty = 1$, donde $D : L_2 \rightarrow L_2$ está definida por $Df = Ff$ para cada $f \in L_2$.

Demostración. Por el corolario 5.3.12, es suficiente verificar la primera parte del lema. T es autoadjunto y acotado, además tenemos:

$$\|T\| = \|P_A P_B P_A\| = \|P_B P_A\|^2 = 1.$$

Fijamos algún $x \in H$ y definimos $y = P_A x$.

Como P_B positivo,

$$\langle Tx, x \rangle = \langle P_A P_B P_A x, x \rangle = \langle P_B P_A x, P_A x \rangle = \langle P_B y, y \rangle \geq 0.$$

Esto prueba que T es positivo en H . □

Al insertar más términos en (λ_k) , podemos suponer que la sucesión en (s_k) definida por $s_k = \lfloor 1/\lambda_k \rfloor$, contiene al conjunto de todos los números naturales.

Además, al aumentar ligeramente algunas de las λ_k podemos suponer que $(1/\lambda_k)$ no contiene números enteros.

Lema 2.2.7. $s_k \lambda_k < 1 \leq (s_k + 1) \lambda_k$ para todo $k \in \mathbb{N}$, $s_1 \leq s_2 \leq \dots$ y cada s_k ocurre solo una cantidad finita de veces.

Demostración. Para cada $n \in \mathbb{N}$, definimos $k_0(n) := \min\{k | s_k = n\}$ y $k_1(n) := \max\{k | s_k = n\}$. Por tanto, $k_1(n) = k_0(n+1) - 1$, $k_0(n) \rightarrow \infty$ y

$$s_{k_0(n)-1} = n - 1 < n = s_{k_0(n)} = s_{k_0(n)+1} = \dots = s_{k_0(n+1)-1} < n + 1 = s_{k_0(n+1)}.$$

Definimos $\alpha_n := (\lambda_{k_0(n)} \cdot n)^{1/2k_1(n)}$ para cada $n \in \mathbb{N}$. □

Lema 2.2.8. Para cada $n \in \mathbb{N}$,

$$1 > \lambda_{k_0(n)} \cdot s_{k_0(n)} = \lambda_{k_0(n)} \cdot n \geq 1 - \lambda_{k_0(n)} \quad (2.6)$$

$$0 < \alpha_n < 1 \text{ y } \alpha_n \rightarrow 1. \quad (2.7)$$

Demostración. Por definición y el lema 2.2.7, tenemos que

$$\lambda_{k_0(n)} \cdot n = \lambda_{k_0(n)} s_{k_0(n)} < 1$$

$$1 \leq \lambda_{k_0(n)} \cdot (s_{k_0(n)} + 1)$$

La última desigualdad implica que

$$1 - \lambda_{k_0(n)} \leq \lambda_{k_0(n)} s_{k_0(n)} = \lambda_{k_0(n)} \cdot n.$$

Además $\lambda_{k_0(n)} \cdot n < 1$ conlleva que $0 < \alpha_n < 1$. Ya que $\lambda_{k_0(n)} \rightarrow 0$, (2.6) implica que $\lambda_{k_0(n)} \cdot n \rightarrow 1$. Esto junto con $k_1(n) \rightarrow \infty$, prueba que $\alpha_n \rightarrow 1$. □

Lema 2.2.9. $\mu\{F^{-1}([1, \infty))\} = 0$.

Demostración. Para ver esto, definamos $S := F^{-1}([1, \infty))$ e $y = U^{-1}(X_S)$ donde X_S denota la función característica de S : $X_S(t) = 1$ si $t \in S$ y $X_S(t) = 0$ en otro caso.

Hay que probar que $\mu(S) = 0$. Ya que

$$\|y\| = \|U^{-1}(X_S)\| = \|X_S\| = \left(\int_S 1 d\mu \right)^{1/2} = [\mu(S)]^{1/2} \quad (2.8)$$

es suficiente probar que $y = 0$. Usando (2.8), tenemos:

$$\begin{aligned} \|Ty\| &= \|U^{-1}DUy\| = \|U^{-1}D(X_S)\| = \|U^{-1}(F_{X_S})\| \\ &= \|F_{X_S}\| = \left(\int_S F^2 d\mu \right)^{1/2} \text{ ge } \left(\int_S 1 d\mu \right)^{1/2} = \|y\|. \end{aligned}$$

Esto demuestra que $\|Ty\| \geq \|y\|$. Pero como $\|T\| = 1$ por el lema 2.2.6, debemos tener $\|Ty\| = \|y\|$. Deducimos que

$$\|y\| = \|Ty\| = \|P_A P_B P_A y\| \leq \|P_B P_A y\| \leq \|P_A y\| \leq \|y\|. \quad (2.9)$$

Por tanto en (2.9) todo debe ser igualdades, entonces $y \in A \cap B = \{0\}$ y luego $y = 0$. \square

Lema 2.2.10. *Si A y B son subespacios cerrados con $A \cap B = \{0\}$, entonces la aplicación no expansiva $P_B P_A$ es estricta, es decir,*

$$\|P_B P_A(x) - P_B P_A(y)\| < \|x - y\|$$

para cada x, y , con $x \neq y$. Como consecuencia de esto, $P_A P_B P_A$ es también estricta. Este lema es consecuencia inmediata del lema anterior.

Lema 2.2.11. *Para cada $\epsilon > 0$, $\mu\{F^{-1}((1 - \epsilon, 1))\} > 0$.*

Demostración. Ya que $F^{-1}\{1\} \subset F^{-1}([1, \infty))$, se deduce que el lema 2.2.9 que $\mu\{F^{-1}(\{1\})\} = 0$. Por lo tanto, $\mu\{F^{-1}((1 - \epsilon, 1))\} = \mu\{F^{-1}((1 - \epsilon, 1])\}$ para cada $\epsilon > 0$. Pero dado que $\|F\|_\infty = 1$, tenemos que $\mu\{F^{-1}((1 - \epsilon, 1))\} > 0$ para cada $\epsilon > 0$. \square

Lema 2.2.12. *Para cada $\epsilon > 0$, existe $\epsilon_1 \in (0, \epsilon)$ tal que $\mu\{F^{-1}((1 - \epsilon, 1 - \epsilon_1))\} > 0$.*

Demostración. Para probar esto, es suficiente usar el lema 2.2.11 y observar que $F^{-1}((1 - \epsilon, 1))$ es la unión numerable de los conjuntos $F^{-1}((1 - \epsilon, 1 - \epsilon/2^k))$. \square

Lema 2.2.13. *Existe una sucesión de números reales $(\beta_n) \subset (0, 1)$ tal que $\alpha_n^2 \leq \beta_n < \beta_{n+1} < 1$ y $\mu\{F^{-1}([\beta_n, \beta_{n+1}))\} > 0$ para todo $n \in \mathbb{N}$.*

Demostración. Probemos esto por inducción. Para $n = 1$, tomamos $\beta_1 = \alpha_1^2$, luego $\beta_1 < 1$. Elegimos ahora $\beta_1 < \beta_2 < \dots < \beta_m < 1$ y $\beta_k \geq \alpha_k^2$ para $k = 1, 2, \dots, m$ y $\mu\{F^{-1}([\beta_k, \beta_{k+1}))\} > 0$ para $k = 1, 2, \dots, m - 1$.

Sea $\epsilon := \min\{1 - \alpha_{m+1}^2, 1 - \beta_m\}$. Luego ya que $\epsilon > 0$ y el lema 2.2.12, implica que existe $\epsilon_1 \in (0, \epsilon)$ tal que $\mu\{F^{-1}([1 - \epsilon, 1 - \epsilon_1))\} > 0$.

Tomamos $\beta_{m+1} := 1 - \epsilon_1$. Entonces $\beta_{m+1} > 1 - \epsilon \geq \beta_m$.

Además, $\beta_{m+1} > 1 - \epsilon \geq \alpha_{m+1}^2$. Finalmente,

$$\mu\{F^{-1}([\beta_m, \beta_{m+1}))\} \geq \mu\{F^{-1}([1 - \epsilon, 1 - \epsilon_1))\} > 0.$$

\square

Definición 2.2.14. Sea β_n el definido en el lema 2.2.13, para cada $n \in \mathbb{N}$, tenemos $S_n := F^{-1}([\beta_n, \beta_{n+1}])$ y se define el vector $e_n \in H$ y su imagen f_n por:

$$e_n = \frac{1}{\sqrt{\mu(S_n)}} U^{-1}(X_{S_n})$$

$$f_n = Ue_n = \frac{1}{\sqrt{\mu(S_n)}} X_{S_n}.$$

Lema 2.2.15. $\|e_n\| = \|f_n\| = 1$ para cada $n \in \mathbb{N}$.

Demostración. Esto se deduce de

$$\|e_n\| = \|Ue_n\| = \|f_n\| = \frac{1}{\sqrt{\mu(S_n)}} \|X_{S_n}\| = 1. \quad (2.10)$$

□

Lema 2.2.16.

1. $T^k = (U^{-1}DU)^k = U^{-1}D^kU$.
2. $D^k f = F^k f$ para todo $f \in L_2(\Omega, \mu)$.
3. Si $f, g \in L_2(\Omega, \mu)$ y $f(t)g(t) = 0$ para μ , para casi todo t , luego $\langle D^j f, D^k g \rangle = 0$ para todo $j, k \in \mathbb{N} \cup \{0\}$.

Lema 2.2.17. Para todos los enteros $j, k \in \mathbb{N} \cup \{0\}$ y $m, n \in \mathbb{N}$ con $m \neq n$, tenemos $\langle T^j e_m, T^k e_n \rangle = 0$. En particular, tomando $j = k = 0$, $\{e_n | n \in \mathbb{N}\}$ es un conjunto ortonormal.

Demostración. Para verificar esto, notemos que (2.10) implica que $f_r = Ue_r = \frac{1}{\sqrt{\mu(S_r)}} X_{S_r}$ para cada $r \in \mathbb{N}$. Luego $X_{S_n} X_{S_m} = X_{S_n \cap S_m}$ ya que $S_n \cap S_m = \emptyset$. Luego $f_n f_m = 0$, usando 1 y 3 del lema 2.2.16, obtenemos que:

$$\langle T^j e_m, T^k e_n \rangle = \langle U^{-1} D^j U e_m, U^{-1} D^k U e_n \rangle = \langle D^j f_m, D^k f_n \rangle = 0.$$

□

Lema 2.2.18. $\beta_{n+1}^k \geq \|T^k e_n\| \geq \beta_n^k \geq \alpha_n^{2k}$ para todo $k, n \in \mathbb{N}$.

Demostración. La última desigualdad se deduce del lema 2.2.13. Ahora observemos que

$$\begin{aligned} \|T^k e_n\|^2 &= \|U^{-1} D^k U e_n\|^2 = \left\| D^k \frac{1}{\sqrt{\mu(S_n)}} X_{S_n} \right\|^2 \\ &= \int [D^k \frac{1}{\sqrt{\mu(S_n)}} X_{S_n}]^2 d\mu \\ &= \frac{1}{\sqrt{\mu(S_n)}} \int_{S_n} F^{2k} d\mu. \end{aligned}$$

Además por la definición de S_n (definición 2.2.14), está claro que

$$\beta_n^{2k} \leq \frac{1}{\sqrt{\mu(S_n)}} \int_{S_n} F^{2k} d\mu \leq \beta_{n+1}^{2k}.$$

Tomando raíces cuadradas, se completa la demostración. \square

Definición 2.2.19. Sea $x_\lambda := \sum_{n=1}^{\infty} \frac{1}{n} e_n$. Como $\sum_{n=1}^{\infty} \frac{1}{n^2} < \infty$ y $\{e_n\}$ es un conjunto ortonormal, x_λ es un elemento bien definido de H .

Lema 2.2.20. $\|T^k x_\lambda\| \geq \alpha_n^{2k}/n$ para todo $n, k \in \mathbb{N}$.

Demostración.

$$\begin{aligned} \|T^k x_\lambda\|^2 &= \langle T^k x_\lambda, T^k x_\lambda \rangle \\ &= \langle T^k \left(\sum_{n=1}^{\infty} e_n/n \right), T^k \left(\sum_{m=1}^{\infty} e_m/m \right) \rangle \\ &= \sum_{n=1}^{\infty} 1/n \sum_{m=1}^{\infty} 1/m \langle T^k e_n, T^k e_m \rangle \\ &= \sum_{n=1}^{\infty} 1/n^2 \|T^k e_n\|^2 \quad (\text{por el lema 2.2.17}) \\ &\geq 1/n^2 \|T^k e_n\|^2 \\ &\geq \alpha_n^{4k}/n^2 \quad (\text{por el lema 2.2.18}). \end{aligned}$$

Luego

$$\|T^k x_\lambda\| \geq \alpha_n^{2k}/n.$$

\square

Lema 2.2.21. $\|(P_B P_A)^k x_\lambda\| \geq \lambda_k$ para cada $k \in \mathbb{N}$.

Demostración. Fijamos un $k \in \mathbb{N}$ y sea $n = S_k$. Usando el lema 2.2.20,

$$\|(P_B P_A)^k x_\lambda\| \geq \|P_A (P_B P_A)^k x_\lambda\| = \|T^k x_\lambda\| \geq \alpha_n^{2k}/n \geq \alpha_n^{2k_1(n)}/n = \lambda_{k_0(n)} \geq \lambda_k.$$

\square

Lema 2.2.22. Para cada $k \in \mathbb{N}$, $(P_{M_2} P_{M_1})^k - P_M = (P_B P_A)^k$.

Demostración. Sea $P_i = P_{M_i}$ y $P = P_2 P_1$, y sabemos que $P_M = P^n P_M$. Además

$$\begin{aligned}
 P^n P_M &= P^n - P^n P^n (I - P_M) = P^n P_{M^\perp} \\
 &= P^n P_{M^\perp}^n && \text{(porque } P_{M^\perp} \text{ es idempotente)} \\
 &= (P P_{M^\perp})^n && \text{(ya que } P \text{ conmuta con } P_{M^\perp}) \\
 &= [(P_2 P_{M^\perp})(P_1 P_{M^\perp})]^n && \text{(porque } P_i \text{ conmuta con } P_{M^\perp} \text{ y } P_{M^\perp} \text{ es idempotente)}.
 \end{aligned}$$

□

Combinando los lemas 2.2.21 y 2.2.22, tenemos que

$$\left\| (P_{M_1} P_{M_2})^k(x_\lambda) - P_M(x_\lambda) \right\| \geq \lambda_k$$

para cada $k \in \mathbb{N}$. Esto completa la demostración de la dicotomía.

□

Capítulo 3

El método para varios subespacios

El método de proyecciones alternantes es un esquema iterativo para encontrar la mejor aproximación a cualquier punto dado en un espacio de Hilbert a partir de un conjunto finito de subespacios cerrados. Este capítulo está basado en el libro [3].

3.1. El algoritmo de Dykstra

En esta sección vamos a probar la convergencia de un algoritmo iterativo que permite calcular las mejores aproximaciones de la intersección de un número finito de conjuntos convexos cerrados, es decir, $\bigcap_{i=1}^r K_i$, reduciendo el problema a encontrar las mejores aproximaciones de cada uno de los K_i .

Sea K_i un subconjunto convexo cerrado de un espacio de Hilbert H y sea $K = \bigcap_{i=1}^r K_i$. Asumimos $K \neq \emptyset$, dado un $x \in H$, queremos calcular $P_K(x)$ (proyección no lineal y no expansiva).

Notación 3.1.1. Fijamos un $x \in H$:

$$\begin{aligned}x_0 &:= x, \quad e_{-(r-1)} = \dots = e_{-1} = e_0 = 0 \\x_n &:= P_{K_{[n]}}(x_{n-1} + e_{n-r}) \quad (n = 1, 2, \dots) \\e_n &:= x_{n-1} + e_{n-r} - x_n = x_{n-1} + e_{n-r} - P_{K_{[n]}}(x_{n-1} + e_{n-r})\end{aligned}$$

donde

$$[n] = \{1, 2, \dots, r\} \cap \{n - kr \mid k = 0, 1, \dots\}. \quad (3.1)$$

Tenemos que probar que

$$\lim_n \|x_n - P_K(x)\| = 0 \quad (3.2)$$

Notación 3.1.2. $P_i := P_{K_i}$ para $i = 1, 2, \dots, r$.

Lema 3.1.3. Para cada n ,

$$\langle x_n - y, e_n \rangle \geq 0 \text{ para todo } y \in K_{[n]}$$

Demostración. Sea $x = x_{n-1} + e_{n-r}$, $y_0 = P_{[n]}(x_{n-1} + e_{n-r})$ y $K = K_{[n]}$, obtenemos para todo $y \in K_{[n]}$

$$\langle x_n - y, e_n \rangle = \langle P_{[n]}(x_{n-1} + e_{n-r}) - y, x_{n-1} + e_{n-r} - P_{[n]}(x_{n-1} + e_{n-r}) \rangle \geq 0$$

□

Lema 3.1.4. Para cada $n \geq 0$

$$x - x_n = e_{n-(r-1)} + e_{n-(r-2)} + \cdots + e_{n-1} + e_n.$$

Demostración. Por inducción en n . Para $n = 0$, $x - x_0 = x - x = 0$, y $e_{n-(r-1)} + e_{n-(r-2)} + \cdots + e_{n-1} + e_n = 0$ por (3.1). Luego, suponemos que el resultado es cierto para n , veamos que también lo es para $n + 1$:

$$\begin{aligned} x - x_{n+1} &= (x - x_n) + (x_n - x_{n+1}) \\ &= (e_{n-(r-1)} + e_{n-(r-2)} + \cdots + e_{n-1} + e_n) + (e_{n+1} - e_{n+1-r}) \\ &= e_{n-(r-2)} + e_{n-(r-3)} + \cdots + e_n + e_{n+1} \\ &= e_{n+1-(r-1)} + e_{n+1-(r-2)} + \cdots + e_n + e_{n+1}. \end{aligned}$$

□

Lema 3.1.5. Para cada $n \in \mathbb{N}$, $0 \leq m \leq n$ e $y \in K$,

$$\begin{aligned} \|x_m - y\|^2 &= \|x_n - y\|^2 + \sum_{k=m+1}^n \|x_k - x_{k-1}\|^2 + 2 \sum_{k=m+1}^n \langle e_{k-r}, x_{k-r} - x_k \rangle \\ &\quad + 2 \sum_{k=n-(r-1)}^n \langle e_k, x_k - y \rangle - 2 \sum_{k=m-(r-1)}^m \langle e_k, x_k - y \rangle. \end{aligned} \quad (3.3)$$

Demostración. Para un conjunto de vectores $\{y_m, y_{m+1}, \dots, y_{n+1}\}$ en H , tenemos

$$\begin{aligned} \|y_m - y_{n+1}\| &= \|(y_m - y_{m+1}) + (y_{m+1} - y_{m+2}) + \cdots + (y_n - y_{n+1})\|^2 \\ &= \sum_{k=m+1}^{n+1} \|y_{k-1} - y_k\|^2 + 2 \sum_{m+1 \leq i < j \leq n+1} \langle y_{i-1} - y_i, y_{j-1} - y_j \rangle \\ &= \|y_m - y_{n+1}\|^2 + \sum_{k=m+1}^n \|y_{k-1} - y_k\|^2 + 2 \sum_{i=m+1}^n \left(\sum_{j=i+1}^{n+1} \langle y_{i-1} - y_i, y_{j-1} - y_j \rangle \right). \end{aligned} \quad (3.4)$$

Pero

$$\begin{aligned} \sum_{i=m+1}^n \left(\sum_{j=i+1}^{n+1} \langle y_{i-1} - y_i, y_{j-1} - y_j \rangle \right) &= \sum_{i=m+1}^n \langle y_{i-1} - y_i, \sum_{j=i+1}^{n+1} (y_{j-1} - y_j) \rangle \\ &= \sum_{i=m+1}^n \langle y_{i-1} - y_i, y_i - y_{n+1} \rangle. \end{aligned} \quad (3.5)$$

Sustituyendo y_i por x_i , para todo $i \leq n$ e $y_{n+1} = y$ en (3.4) obtenemos

$$\begin{aligned} \sum_{i=m+1}^n \langle y_{i-1} - y_i, y_i - y_{n+1} \rangle &= \sum_{i=m+1}^n \langle x_{i-1} - x_i, x_i - y \rangle \\ &= \sum_{i=m+1}^n \langle e_i - e_{i-r}, x_i - y \rangle \\ &= \sum_{i=m+1}^n \langle e_i, x_i - y \rangle - \sum_{i=m+1}^n \langle e_{i-r}, x_i - y \rangle \\ &= \sum_{i=m+1}^n \langle e_i, x_i - y \rangle - \sum_{i=m+1}^n [\langle e_{i-r}, x_i - x_{i-r} \rangle + \langle e_{i-r}, x_{i-r} - y \rangle] \\ &= \sum_{i=m+1}^n \langle e_i, x_i - y \rangle - \sum_{i=m+1}^n \langle e_{i-r}, x_{i-r} - y \rangle + \sum_{i=m+1}^n \langle e_{i-r}, x_{i-r} - x_i \rangle \\ &= \sum_{i=m+1}^n \langle e_i, x_i - y \rangle - \sum_{i=m+1-r}^{n-r} \langle e_i, x_i - y \rangle + \sum_{i=m+1}^n \langle e_{i-r}, x_{i-r} - x_i \rangle \\ &= \sum_{i=m+1}^n \langle e_i, x_i - y \rangle - \sum_{i=m-r+1}^m \langle e_i, x_i - y \rangle + \sum_{i=m+1}^n \langle e_{i-r}, x_{i-r} - x_i \rangle. \end{aligned}$$

Volviendo a sustituir estos mismos valores por y_i en (3.4) y usando la relación anterior, tenemos que

$$\begin{aligned} \|x_m - y\|^2 &= \|x_n - y\|^2 + \sum_{k=m+1}^n \|x_{k-1} - x_k\|^2 + 2 \sum_{i=n-r+1}^n \langle e_i, x_i - y \rangle \\ &\quad + 2 \sum_{i=m+1}^n \langle e_{i-r}, x_{i-r} - x_i \rangle - 2 \sum_{i=m-r+1}^m \langle e_i, x_i - y \rangle \end{aligned}$$

lo que verifica (3.3). □

Lema 3.1.6. $\{x_n\}$ es una sucesión acotada y

$$\sum_1^{\infty} \|x_{k-1} - x_k\|^2 < \infty. \quad (3.6)$$

En particular tenemos

$$\|x_{n-1} - x_n\| \rightarrow 0 \text{ cuando } n \rightarrow \infty. \quad (3.7)$$

Demostración. Estableciendo $m = 0$ en (3.3), el tercer y el cuarto término son no negativos, por el lema 3.1.3, mientras que el último término es cero. Así tenemos que

$$\|x_0 - y\|^2 \leq \|x_n - y\|^2 + \sum_{k=1}^n \|x_k - x_{k-1}\|^2$$

para todo n , por lo que $\{\|x_n - y\|\}$ acotado y se cumple (3.6). Por lo tanto, x_n es acotado también. Finalmente, (3.6) implica (3.7). \square

Lema 3.1.7. Para cualquier $n \in \mathbb{N}$

$$\|e_n\| \leq \sum_{k=1}^n \|x_{k-1} - x_k\|. \quad (3.8)$$

Demostración. Por inducción en n .

Para $n=1$,

$$\|e_1\| = \|x_0 - x_1 - e_{1-r}\| = \|x_0 - x_1\|$$

ya que $e_i = 0$ para todo $i \leq 0$. Ahora asumimos que (3.6) es cierto para todo $m \leq n$. Luego

$$\begin{aligned} \|e_{n+1}\| &= \|x_n - x_{n+1} + e_{n+1-r}\| \leq \|x_n - x_{n+1}\| + \|e_{n+1-r}\| \\ &\leq \|x_n - x_{n+1}\| + \sum_{k=1}^{n+1-r} \|x_{k-1} - x_k\| \leq \sum_{k=1}^{n+1} \|x_{k-1} - x_k\| \end{aligned}$$

implica que (3.8) es cierto para $n + 1$. \square

Lema 3.1.8.

$$\liminf_n \sqrt{n} \sum_{k=n-(r-1)}^n |\langle x_k - x_n, e_k \rangle| = 0.$$

Demostración. Usando la desigualdad de Schwarz y el lema 3.1.7, tenemos

$$\begin{aligned}
\sum_{k=n-(r-1)}^n |\langle x_k - x_n, e_k \rangle| &\leq \sum_{k=n-(r-1)}^n \|e_k\| \|x_k - x_n\| \\
&\leq \sum_{k=n-(r-1)}^n \left(\sum_{j=1}^k \|x_{j-1} - x_j\| \right) \left(\sum_{i=k+1}^n \|x_{i-1} - x_i\| \right) \\
&\leq r \sum_{j=1}^n \|x_{j-1} - x_j\| \left(\sum_{i=n-(r-2)}^n \|x_{i-1} - x_i\| \right).
\end{aligned}$$

Definiendo $a_i = \|x_{i-1} - x_i\|$, es suficiente probar que

$$\liminf_n \sum_{j=1}^n a_j \left(\sum_{i=n-(r-2)}^n a_i \right) = 0. \quad (3.9)$$

Pero

$$A := \sum_1^\infty a_i^2 = \sum_1^\infty \|x_{i-1} - x_i\|^2 < \infty$$

por el lema 3.1.6. Usando de nuevo la desigualdad de Cauchy-Schwarz, obtenemos

$$\sum_1^n a_j \leq \sqrt{n} \left(\sum_1^n a_j^2 \right)^{1/2} \leq \sqrt{n} A^{1/2}$$

para cada n . Para verificar (3.9), basta comprobar que

$$\liminf_n \sqrt{n} \sum_{k=n-(r-2)}^n a_i = 0. \quad (3.10)$$

Para probar esto, definimos

$$\rho := \liminf_n \sqrt{n} \sum_{k=n-(r-2)}^n a_i.$$

Si (3.10) fuese falso, tendríamos que $\rho > 0$. Asumimos $\rho < \infty$ (para $p = \infty$ la demostración es similar). Luego

$$\sqrt{n} \sum_{k=n-(r-2)}^n a_i > \frac{1}{2} \rho \quad (\text{para } n \text{ suficientemente grande}).$$

Así $\sum_{i=n-(r-2)}^n a_i > \rho/(2\sqrt{n})$ para n suficientemente grande. Usando la desigualdad de Schwarz, se deduce que

$$\frac{\rho^2}{4n} < \left(\sum_{i=n-(r-2)}^n a_i \right)^2 \leq (r-1) \sum_{i=n-(r-2)}^n a_i^2 \quad (\text{para } n \text{ suficientemente grande})$$

Para algún entero N ,

$$\frac{\rho^2}{4n} < (r-1) \sum_{i=n-(r-2)}^n a_i^2 \quad \text{para todo } n \geq N.$$

Por lo tanto,

$$\begin{aligned} \infty &= \sum_{n=N}^{\infty} \frac{\rho^2}{4n} \leq (r-1) \sum_{n=N}^{\infty} \sum_{i=n-(r-2)}^n a_i^2 \\ &\leq (r-1) \sum_{n=N}^{\infty} [a_{n-(r-2)}^2 + a_{n-(r-3)}^2 + \cdots + a_n^2] \\ &\leq (r-1)^2 \sum_{n=N}^{\infty} a_i^2 < \infty. \end{aligned}$$

□

Lema 3.1.9. *Existe una subsucesión $\{x_{n_j}\}$ de $\{x_n\}$ tal que*

$$\limsup_j \langle y - x_{n_j}, x - x_{n_j} \rangle \leq 0 \quad \text{para cada } y \in K \text{ y} \quad (3.11)$$

$$\lim_j \sum_{k=n_j-(r-1)}^{n_j} |\langle x_k - x_{n_j}, e_k \rangle| = 0. \quad (3.12)$$

Demostración. Usando el lema 3.1.4, tenemos que para todo $y \in K$, $n \geq 0$

$$\begin{aligned} \langle y - x_n, x - x_n \rangle &= \langle y - x_n, e_{n-(r-1)} + e_{n-(r-2)} + \cdots + e_n \rangle \\ &= \sum_{k=n-(r-1)}^n \langle y - x_n, e_k \rangle \\ &= \sum_{k=n-(r-1)}^n \langle y - x_k, e_k \rangle + \sum_{k=n-(r-1)}^n \langle x_k - x_n, e_k \rangle. \end{aligned} \quad (3.13)$$

Por el lema 3.1.3, el primer sumando no es mayor que cero. Por lo tanto,

$$\langle y - x_n, x - x_n \rangle \leq \sum_{k=n-(r-1)}^n \langle x_k - x_n, e_k \rangle. \quad (3.14)$$

Deducimos, por el lema 3.1.8, que la subsucesión $\{n_j\}$ de \mathbb{N} tal que

$$\lim_j \sum_{k=n_j-(r-1)}^{n_j} |\langle x_k - x_{n_j}, e_k \rangle| = 0. \quad (3.15)$$

La parte derecha de (3.14) no depende de y , de aquí se deduce que (3.15) implica que (3.11) y (3.12) son ciertos. \square

Recordatorio 3.1.10. Sea K_i un subconjunto convexo cerrado de un espacio de Hilbert H y sea $K = \bigcap_{i=1}^r K_i$. Para cada $x \in H$:

$$\begin{aligned} x_0 &:= x, \quad e_{-(r-1)} = \dots = e_{-1} = e_0 = 0 \\ x_n &:= P_{K[n]}(x_{n-1} + e_{n-r}) \quad (n = 1, 2, \dots) \\ e_n &:= x_{n-1} + e_{n-r} - x_n = x_{n-1} + e_{n-r} - P_{K[n]}(x_{n-1} + e_{n-r}) \end{aligned}$$

donde

$$[n] = \{1, 2, \dots, r\} \cap \{n - kr \mid k = 0, 1, \dots\}. \quad (3.16)$$

Teorema 3.1.11 (Teorema de Boyle-Dykstra). *Sea K_1, \dots, K_r subconjuntos convexos cerrados del espacio de Hilbert H tal que $K = \bigcap_{i=1}^r K_i \neq \emptyset$. Para cada $x \in H$, definimos la sucesión $\{x_n\}$ como en (3.16). Luego*

$$\lim_n \|x_n - P_K(x)\| = 0. \quad (3.17)$$

Demostración. Por el lema 3.1.9, existe una subsucesión $\{x_{n_j}\}$ tal que

$$\limsup_j \langle y - x_{n_j}, x - x_{n_j} \rangle \leq 0 \quad \text{para cada } y \in K. \quad (3.18)$$

Como x_{n_j} es acotado por el lema 3.1.6, hay un $y_0 \in H$ tal que

$$x_{n_j} \rightarrow y_0 \quad (3.19)$$

y

$$\lim_j \|x_{n_j}\| \text{ existe.} \quad (3.20)$$

Tenemos que

$$\|y_0\| \leq \liminf_j \|x_{n_j}\| = \lim_j \|x_{n_j}\|. \quad (3.21)$$

Como hay un número finito de conjuntos K_i , un infinito número de los x_{n_j} debe estar en un solo conjunto K_{i_0} . Por (3.7), $x_n - x_{n-1} \rightarrow 0$. Por una aplicación repetida de este hecho, observamos que todas las sucesiones $\{x_{n_j+1}\}, \{x_{n_j+2}\}, \dots$ convergen débilmente a y_0 y por lo tanto, $y_0 \in K_i$ para todo i . Esto significa que

$$y_0 \in K. \quad (3.22)$$

Para cualquier $y \in K$, (3.21) y (3.18) implica que

$$\begin{aligned} \langle y - y_0, x - y_0 \rangle &= \langle y, x \rangle - \langle y, y_0 \rangle - \langle y_0, x \rangle + \|y_0\|^2 \\ &\leq \lim_j [\langle y, x \rangle - \langle y, x_{n_j} \rangle - \langle x_{n_j}, x \rangle + \|x_{n_j}\|^2] \\ &= \lim_j \langle y - x_{n_j}, x - x_{n_j} \rangle \leq 0. \end{aligned}$$

Por el teorema de mejor aproximación de conjuntos convexos,

$$y_0 = P_K(x). \quad (3.23)$$

Además, poniendo $y = y_0$ en las anteriores desigualdades obtenemos la igualdad y por lo tanto,

$$\lim_j \|x_{n_j}\|^2 = \|y_0\|^2 \quad (3.24)$$

y

$$\lim_j \langle y_0 - x_{n_j}, x - x_{n_j} \rangle = 0. \quad (3.25)$$

Por (3.19) y (3.24), se tiene que

$$\|x_{n_j} - y_0\| \rightarrow 0. \quad (3.26)$$

Por lo tanto,

$$\|x_{n_j} - P_K(x)\| = \|x_{n_j} - y_0\| \rightarrow 0. \quad (3.27)$$

Para acabar con la demostración, debemos probar que toda la sucesión $\{x_n\}$ converge a y_0 . De la ecuación (3.13) con $y = y_0$ y $n = n_j$, conseguimos

$$\langle y_0 - x_{n_j}, x - x_{n_j} \rangle = \sum_{k=n_j-(r-1)}^{n_j} \langle y_0 - x_k, e_k \rangle + \sum_{k=n_j-(r-1)}^{n_j} \langle x_k - x_{n_j}, e_k \rangle \quad (3.28)$$

El lado izquierdo de (3.28) tiende a cero cuando n tiende a infinito por (3.25), mientras que el segundo sumando de la derecha tiende a cero por (3.14). Por lo tanto,

$$\lim_j \sum_{k=n_j-(r-1)}^{n_j} \langle y_0 - x_k, e_k \rangle = 0. \quad (3.29)$$

Usando los lemas 3.1.5 y 3.1.3 con $m = n_j$ y $y = y_0$, podemos observar que para todo $n \geq n_j$,

$$\begin{aligned} \|x_{n_j} - y_0\|^2 &\geq \|x_n - y_0\|^2 - 2 \sum_{k=n_j-(r-1)}^{n_j} \langle e_k, x_k - y_0 \rangle \\ &\quad \acute{o} \\ \|x_n - y_0\|^2 &\leq \|x_{n_j} - y_0\|^2 + 2 \sum_{k=n_j-(r-1)}^{n_j} \langle e_k, x_k - y_0 \rangle. \end{aligned}$$

Pero ambos términos de la derecha de la desigualdad tienden a 0 cuando $j \rightarrow \infty$ por (3.26) y (3.29). De aquí se deduce que $\lim_n \|x_n - y_0\| = 0$, lo que completa la demostración. \square

El resultado obtenido por von Neumann fue extendido posteriormente por Halperin a una familia finita de subespacios cerrados.

Corolario 3.1.12 (Halperin). *Sea M_1, \dots, M_r subespacios cerrados en un espacio de Hilbert H y $M = \bigcap_{i=1}^r M_i$, luego para cada $x \in H$ se tiene que*

$$\lim_{n \rightarrow \infty} (P_{M_r} \cdots P_{M_1})^n(x) = P_M(x).$$

Este teorema ya está demostrado para subconjuntos convexos cerrados, por lo que también es cierto para subespacios cerrados.

3.2. Convergencia geométrica

Si $M_1 \dots M_r$ son subespacios cerrados en un espacio de Hilbert H y $M = \bigcap_{i=1}^r M_i$, por el corolario anterior tenemos que para todo $x \in H$:

$$\lim_{n \rightarrow \infty} \|(P_{M_r} \cdots P_{M_1})^n(x) - P_M(x)\| = 0$$

Entonces

$$\|(P_{M_r} \cdots P_{M_1})^n(x) - P_M(x)\| \leq B_r(n)\|x\| \quad (3.30)$$

donde $B_r(n) := \|(P_{M_r} \cdots P_{M_1})^n - P_M\|$ siendo esta la cota más pequeña para (3.30) independientemente de x .

Lema 3.2.1. *Sea $M_1 \cdots M_r$ subespacios cerrados de H y $M = \bigcap_{i=1}^r M_i$. Entonces:*

- (1) P_{M^\perp} es idempotente y autoadjunto.
- (2) P_{M^\perp} conmuta con cada $P_{M_i^\perp}$ y con $P_{M_r} \cdots P_{M_1}$.
- (3) Para $i = 1, \dots, r$, $P_{M_i}(x) = x$ para todo $x \in M$ y $P_{M_i}(M^\perp) \in M^\perp$. Por lo tanto, $P_{M_r} \cdots P_{M_1}x = x$, para todo $x \in M$ y $P_{M_r} \cdots P_{M_1}(M^\perp) \subset M^\perp$.

Demostración.

(1) Probemos primero que P_M es idempotente, es decir, que $P_M(P_M(x)) = P_M(x)$, o lo que es lo mismo que $P_M^2 = P_M$. Esto se debe a que cada $y \in K$ es su propia mejor aproximación en M : $P_M(y) = y$. Como P_M es idempotente, entonces P_{M^\perp} también lo es. Veamos ahora que es autoadjunto, para todo $x, y \in H$, la caracterización de la mejor aproximación de subespacios implica que $\langle P_M(x), y - P_M(x) \rangle = 0$ y por lo tanto,

$$\langle P_M(x), y \rangle = \langle P_M(x), P_M(y) \rangle. \quad (3.31)$$

Cambiando x por y y viceversa en (3.31) obtenemos

$$\langle x, P_M(y) \rangle = \langle P_M(y), x \rangle = \langle P_M(y), P_M(x) \rangle = \langle P_M(x), P_M(y) \rangle = \langle P_M(x), y \rangle.$$

Hemos probado que P_M es autoadjunto, por lo que P_{M^\perp} también lo es.

(2) $M \in M_i$ para todo i ya que P_M conmuta con P_i , además tenemos que

$$P_{M^\perp}P_i = (I - P_M)P_i = P_i - P_MP_i = P_i - P_iP_M = P_i(I - P_M) = P_iP_{M^\perp}.$$

Luego P_{M^\perp} conmuta con P_i y además P_{M^\perp} conmuta con $P_r \dots P_1$.

(3) Si $x \in M$, entonces $x \in M_i$ para todo i , por lo que $P_ix = x$. Como P_{M^\perp} y P_{M_i} conmutan, para todo $x \in M^\perp$ se tiene:

$$P_{M_i}(x) = P_{M_i}P_{M^\perp}(x) = P_{M^\perp}P_{M_i} \in M^\perp$$

y $P_{M_i} \subset M^\perp$. Además $P_{M_r} \dots P_{M_1}x = x$ y $P_{M_r} \dots P_{M_1}(M^\perp) \subset M^\perp$. □

Lema 3.2.2. Sea $M_1 \dots M_r$ subespacios cerrados en H y $M = \bigcap_{i=1}^r M_i$. Para todo $n \in \mathbb{N}$,

$$(P_{M_r} \dots P_{M_1})^n - P_M = (P_{M_r} \dots P_{M_1})^n P_{M^\perp} = (P_{M_r} \dots P_{M_1} P_{M^\perp})^n = (Q_r \dots Q_1)^n \quad (3.32)$$

donde $Q_j = P_{M_j \cap M^\perp}$ para $j = 1, \dots, r$. En particular,

$$\|(P_{M_r} \dots P_{M_1})^n - P_{\bigcap_{i=1}^r M_i}\| = \|(Q_r \dots Q_1)^n\| = \|(Q_1 \dots Q_r)^n\|. \quad (3.33)$$

Demostración. Sea $P_i = P_{M_i}$ y $P = P_r \dots P_1$, como el adjunto de $(Q_r \dots Q_1)^n$ es $(Q_1 \dots Q_r)^n$, tenemos la última igualdad de (3.33), por lo que basta probar las otras desigualdades de (3.32). Usando el lema 3.2.1, tenemos que $P_M = P^n P_M$ y que

$$\begin{aligned} P^n - P_M &= P^n - P^n P_M = P^n(I - P_M) = P^n P_{M^\perp} \\ &= P^n P_{M^\perp}^n \quad (P_{M^\perp} \text{ idempotente}) \\ &= (P P_{M^\perp})^n \quad (P \text{ conmuta con } P_{M^\perp}) \\ &= [(P_r P_{M^\perp}) \dots (P_1 P_{M^\perp})]^n \end{aligned}$$

ya que P_i conmuta con P_{M^\perp} y P_{M^\perp} es idempotente. Como hemos demostrado en el lema 3.2.1 (3) que $P_{M_i}(M^\perp) \subset M^\perp$ tenemos que $P_{M_i}P_{M^\perp} = P_{M_i \cap M^\perp} = Q_i$. □

Ahora supongamos $r = 2$, para demostrar que el límite para el teorema 2.2.2 es el más pequeño posible.

Teorema 3.2.3. *Para todo $n \in \mathbb{N}$,*

$$\|(P_{M_2}P_{M_1})^n - P_{M_1 \cap M_2}\| = c(M_1, M_2)^{2n-1}. \quad (3.34)$$

Demostración. Por el lema 3.2.2, tenemos que

$$\|(P_{M_2}P_{M_1})^n - P_{M_1 \cap M_2}\| = \|(Q_1Q_2)^n\| \quad (3.35)$$

con $Q_i = P_{M_i \cap (M_1 \cap M_2)^\perp}$. Como el adjunto de $(Q_1Q_2)^n$ es $(Q_2Q_1)^n$, por tanto

$$\|(Q_1Q_2)^n\|^2 = \|(Q_1Q_2)^n(Q_2Q_1)^n\|$$

y obtenemos de aquí

$$(Q_1Q_2)^n(Q_2Q_1)^n = (Q_1Q_2Q_1)^{2n-1}. \quad (3.36)$$

Como $Q_1Q_2Q_1$ autoadjunto por (3.36) y

$$\|(Q_1Q_2)^n\|^2 = \|(Q_1Q_2Q_1)^{2n-1}\| = \|Q_1Q_2Q_1\|^{2n-1} \quad (3.37)$$

entonces

$$\begin{aligned} \|Q_1Q_2Q_1\| &= \|(Q_1Q_2)(Q_2Q_1)\| = \|(Q_1Q_2)(Q_1Q_2)^*\| = \|Q_1Q_2\|^2 \\ &= \|P_{M_1 \cap (M_1 \cap M_2)^\perp}P_{M_2 \cap (M_1 \cap M_2)^\perp}\|^2 = c(M_1, M_2)^2. \end{aligned}$$

Por (3.8), $\|(Q_1Q_2)^n\|^2 = c(M_1, M_2)^{2(2n-1)}$ y sustituyendo en (3.35) tenemos el resultado. \square

Observación 3.2.4. Podemos obtener una cota de $B_r(n)$ del lema 3.2.2

$$B_r(n) = \|(P_{M_r} \cdots P_{M_1})^n - P_M\| \leq \|P_{M_r} \cdots P_{M_1}P_{M^\perp}\|^n$$

para todo n , donde $M = \bigcap_{i=1}^r M_i$.

En particular,

$$B_2(n) = \|(Q_1Q_2)^n\| \leq \|Q_1Q_2\|^n = c(M_1, M_2)^n.$$

Capítulo 4

Diversas aplicaciones del método

En este capítulo estudiaremos diferentes áreas de las matemáticas donde el método de las proyecciones alternantes ha desempeñado un papel importante.

4.1. Sistemas de ecuaciones lineales

En primer lugar, veremos que este método nos permite encontrar una solución a los sistemas de ecuaciones lineales, este método iterativo es llamado el método de Kaczmarz [3].

Consideremos el típico problema de resolución de sistema de ecuaciones lineales

$$Ax = b$$

donde la matriz $A \in \mathbb{R}^{r \times n}$, $x \in \mathbb{R}^n$ y $b \in \mathbb{R}^r$. Este problema puede generalizarse para cualquier espacio de Hilbert H . Encontrar una solución de $Ax = b$ es equivalente a encontrarla en la intersección de los r hiperplanos cerrados.

Sea H un espacio de Hilbert, $y_1 \dots y_n \in H \setminus \{0\}$, $c_1, \dots, c_r \in \mathbb{R}$, queremos encontrar un punto $x \in H$, si existe, que satisface las ecuaciones:

$$\langle x, y_i \rangle = c_i \quad (i = 1, 2, \dots, r).$$

Sea $K_i := \{y \in H \mid \langle y, y_i \rangle = c_i\}$ y $K = \bigcap_{i=1}^r K_i$. Además x satisface las ecuaciones si y solo si $x \in K$.

Suponiendo que $K \neq \emptyset$, la proyección sobre los hiperplanos cerrados K_i viene dado por la siguiente expresión

$$P_{K_i}(z) = z - \frac{y_i}{\|y_i\|} \cdot [\langle z, y_i \rangle - c_i].$$

Para encontrar dicha solución, fijamos un $x_0 \in \mathbb{R}^r$ y por inducción definimos

$$x_n = (P_{K_r} P_{K_{r-1}} \cdots P_{K_1}) x_{n-1}.$$

Entonces, el método implica que $x_n \rightarrow P_K x_0$ cuando $n \rightarrow \infty$. Por tanto, $P_K x_0$ es la solución mas cercana a x_0 del conjunto de ecuaciones anteriores.

En particular, si el espacio de Hilbert es $l_2(N)$ y $y_i = (a_{i1}, \dots, a_{iN})$ para $i = 1, \dots, r$, las ecuaciones se escriben como

$$\sum_{j=1}^N a_{ij} x(j) = c_i. \quad (4.1)$$

Comenzando con $x_0 = 0$, el algoritmo de von Neumann produce una sucesión de puntos $\{x_n\}$ en $l_2(N)$ que converge geoméricamente a la única solución de (4.1).

Muchos otros autores han dado extensiones y generalizaciones del método de Kaczmarz.

4.2. Sistema de inecuaciones lineales

Sea H un espacio de Hilbert, $y_1, \dots, y_r \in H \setminus \{0\}$, $c_1, \dots, c_r \in \mathbb{R}$, queremos encontrar un punto $x \in H$, si existe, que satisface las inecuaciones:

$$\langle x, y_i \rangle \leq c_i \quad (i = 1, 2, \dots, r). \quad (4.2)$$

Sea $K_i := \{y \in H \mid \langle y, y_i \rangle \leq c_i\}$ y $K = \bigcap_{i=1}^r K_i$. x resuelve (4.2) si y solo si $x \in K$.

Asumimos $K \neq \emptyset$. Para obtener un punto en K , comenzamos con cualquier $x_0 \in H$ y generamos la sucesión $\{x_n\}$ según el algoritmo de Dykstra, que converge a $P_K(x_0)$.

Veamos el caso particular $H = l_2(N)$ y $y_i = (a_{i1}, \dots, a_{iN})$ para $i = 1, \dots, r$, las inecuaciones se escriben como

$$\sum_{j=1}^N a_{ij} x(j) \leq c_i. \quad (4.3)$$

Por lo que hemos observado, el algoritmo de Dykstra es una forma eficaz de resolver sistema de inecuaciones lineales.

4.3. Restauración de imágenes

René Escalante [4] describe el proceso de restauración de imágenes usando el método de las proyecciones alternantes donde establece un esquema iterativo para resolver dicho problema.

Para ello, consideramos dos subespacios cerrados A y B que están contenidos en un espacio de Hilbert H y supondremos que un elemento $x \in H$ pertenece también al subespacio B ,

del cual sólo se conoce su proyección sobre el subespacio cerrado A , es decir, $y = P_A x$. Por medio de un proceso inverso se reconstruye la imagen x a partir de y , es decir, se desea encontrar la mejor aproximación a la imagen original. Todo elemento x en un espacio de Hilbert H se puede escribir como $x = z + w$, donde $z \in A$ y $w \in A^\perp$ y $P_A = (I - P_{A^\perp})$. Además como $x \in B$, $x = P_B x$. Tenemos

$$\begin{aligned} y &= P_A x \\ &= P_A P_B x \\ &= (I - P_{A^\perp}) P_B x \\ &= P_B x - P_{A^\perp} P_B x \\ &= x - P_{A^\perp} P_B x. \end{aligned}$$

Por lo que, $x = y + P_{A^\perp} P_B x$. Entonces tenemos el siguiente esquema iterativo:

1. $x_{n+1} = y + P_{A^\perp} P_B x_n$.
2. $y = x - P_{A^\perp} P_B x$.
3. $\lim_{n \rightarrow \infty} (P_A P_B)^n x = P_{A \cap B} x$.
4. $x_1 = y$.

Aplicando 1 de forma sucesiva y suponiendo 2 obtenemos

$$x_{n+1} = x - (P_{A^\perp} P_B)^n x + (P_{A^\perp} P_B)^n x_1.$$

Y por 3 y 4, llegamos a

$$\lim_{n \rightarrow \infty} x_{n+1} = x - P_{A^\perp \cap B} x + P_{A^\perp \cap B} y.$$

Este esquema converge a x , solución del problema, si $A^\perp \cap B = \{0\}$ o equivalentemente si $\|P_{A^\perp} P_B\| < 1$. Si ninguna de estas dos condiciones se cumple no podemos asegurar la convergencia del esquema iterativo.

Para las siguientes aplicaciones nos hemos basado en el libro [2].

4.4. Solución al problema de Dirichlet

Schwarz describió lo que él llamo el "Método alternante" para solucionar el problema de Dirichlet en una región irregular en el plano, que es la unión de regiones regulares. Este método permite obtener una solución de dicho problema para una ecuación diferencial de tipo elíptica en los dominios D , que se puede representar como la unión de un número finito de dominios D_j ($j = 1, \dots, r$) en los que ya se conoce la solución al problema de Dirichlet.

Por ejemplo, supongamos D_1 y D_2 dos discos superpuestos en el plano y sea $D = D_1 \cup D_2$. Llamamos ∂S a la frontera de cualquier conjunto S en el plano y supongamos que f es una función real y continua en ∂D . El problema de Dirichlet en D consiste en determinar una función u en D tal que

$$\Delta u = 0 \quad \text{en } D$$

$$u = f \quad \text{sobre } \partial D.$$

En este caso $\Delta u := \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2}$ es el laplaciano de u .

Como $|f| < M$, podemos extender f a una función continua f_1 sobre ∂D_1 . Primero resolvemos el problema de Dirichlet en D_1 con f_1 sobre ∂D_1 . Luego utilizamos esta solución, u_1 , para obtener una extensión continua f_2 de f sobre ∂D_2 . Resolvemos ahora el problema de Dirichlet en D_2 con valor de frontera f_2 sobre ∂D_2 , para obtener una solución armónica v_1 . Continuamos de la misma forma, alternando D_1 y D_2 para obtener una sucesión $u_1, v_1, u_2, v_2 \dots$ tal que $\{u_n\}$ converge uniformemente a una función armónica u sobre D_1 , $\{v_n\}$ converge uniformemente a una función v sobre D_2 y $u = v$ en $D_1 \cap D_2$. Luego la función

$$w = \begin{cases} u & \text{sobre } D_1 \\ v & \text{sobre } D \setminus D_1 \end{cases}$$

resuelve el problema de Dirichlet sobre D .

El método alternante de Schwarz también se utiliza para resolver problemas de valor frontera de naturaleza más general para ecuaciones generales de tipo elíptico (incluidas ecuaciones de un orden mayor que dos) bajo ciertas condiciones adicionales y también en dominios en el espacio.

4.5. Núcleos de Bergman

En 1983 Skwarczynski demostró cómo calcular el núcleo de Bergman para los espacios de Hilbert $L_2(D)$ utilizando este método.

Supongamos que D es una unión de dominios D_i en \mathbb{C}^n con $i = 1, \dots, m$, es decir, $D = \bigcup_1^m D_i$. Suponemos que los núcleos de Bergman K_{D_i} para las regiones D_i son conocidos. Lo que tenemos que hallar es el núcleo de Bergman K_D para la región D .

Sea

$$F_i = \{f \in L_2(D) \mid f \text{ es holomorfa en } D_i\}$$

($i = 1, \dots, m$), por lo que F_i es un subespacio cerrado en $L_2(D)$. Fijando cualquier $t \in D$, imponemos que $t \in D_1$ y definimos $f \in L_2(D)$ como

$$f(z) = \begin{cases} K_{D_1}(z, t) & \text{si } z \in D_1 \\ 0 & \text{si } z \in D \setminus D_1. \end{cases}$$

Luego obtenemos la sucesión

$$f_1 = P_{F_1} f, \quad f_2 = P_{F_2} f_1, \dots, \quad f_m = P_{F_m} f_{m-1},$$

$$f_{m+1} = P_{F_1} f_m, \quad f_{m+2} = P_{F_2} f_{m+1}, \dots$$

que converge en $L_2(D)$ a $K_D(\cdot, t)$.

4.6. Otras aplicaciones

Además de las nombradas anteriormente, el método se ha utilizado para muchas más aplicaciones en diversos campos de las matemáticas. En probabilidad y estadística, por ejemplo, Wiener y Masani han utilizado el método en la teoría lineal de la predicción y Burkholder y Chow estudiaron cuando la convergencia de la norma en $L^2(\mu)$ para el método podría ser sustituido por la convergencia en casi todo punto. Dykstra demostró que el método se mantiene en el espacio $\ell_2(n)$ cuando los subespacios son conos convexos cerrados. También Deutsch lo utilizó para la aproximación de funciones Multivariables, además de ser utilizado para otros estudios como la regresión isótoma y convexa.

Capítulo 5

Apéndice

El objetivo de este capítulo es presentar algunas definiciones, teoremas y proposiciones importantes que serán de gran utilidad para los capítulos anteriores.

5.1. Resultados auxiliares conocidos del grado

Definición 5.1.1. Sea H un espacio vectorial sobre el cuerpo \mathbb{C} de los números complejos, un producto escalar sobre H es una aplicación $\langle \cdot, \cdot \rangle$ de $H \times H$ en \mathbb{C} tal que

$$\langle x, x \rangle \geq 0 \quad \forall x \in H;$$

$$\langle x, x \rangle = 0 \text{ si y solo si } x = 0;$$

$$\langle x, y \rangle = \overline{\langle y, x \rangle} \quad \forall x, y \in H;$$

$$\langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle \quad \forall x, y, z \in H, \alpha, \beta \in \mathbb{C}.$$

Un espacio vectorial dotado de un producto escalar es un espacio prehilbertiano.

Definición 5.1.2. La norma de $x \in H$ es $\|x\| := \sqrt{\langle x, x \rangle}$, se cumple que

$$\|x\| \geq 0, \quad \|x\| = 0 \text{ si y solo si } x = 0;$$

$$\|x + y\| \leq \|x\| + \|y\| \text{ para } x, y \in H;$$

$$\|\alpha x\| = |\alpha| \|x\| \text{ para } x \in H \text{ y } \alpha \in \mathbb{C}.$$

Definiendo $d(x, y) = \|x - y\|$ se obtiene una distancia en H . Un espacio vectorial dotado de una norma es un espacio normado.

Definición 5.1.3. Sea M un espacio métrico, decimos que M es completo cuando toda sucesión de Cauchy converge a un punto de M .

Proposición 5.1.4 (Desigualdad de Cauchy-Schwarz).

$$|\langle x, y \rangle| \leq \|x\| \cdot \|y\| \quad x, y \in H.$$

Definición 5.1.5. Un espacio prehilbertiano que es completo para la distancia inducida por el producto escalar es un espacio de Hilbert.

Definición 5.1.6. Se llama espacio de Banach a un espacio normado que es completo para la distancia inducida por su norma.

Definición 5.1.7. Sea H un espacio prehilbertiano, dos vectores $x, y \in H$ son ortogonales si $\langle x, y \rangle = 0$; se denota por $x \perp y$. Se llama ortogonal de un conjunto A al conjunto A^\perp de todos los vectores ortogonales a los vectores de A , es decir, $A^\perp = \{x \in H : x \perp y, \forall y \in A\}$. A^\perp es un subespacio vectorial cerrado.

Definición 5.1.8. Un subconjunto C de un espacio vectorial es convexo si para todo par de puntos $x, y \in C$ y todo escalar $\lambda \in [0, 1]$, se tiene que $\lambda x + (1 - \lambda)y \in C$.

Teorema 5.1.9 (Teorema de la proyección). *Sea H un espacio de Hilbert y M un subespacio vectorial cerrado de H . Para todo $x \in H$ existe un único par de aplicaciones $P : H \rightarrow M$, $Q : H \rightarrow M^\perp$, tales que $x = Px + Qx$. Las aplicaciones P y Q son las llamadas proyecciones ortogonales de H sobre M y M^\perp , respectivamente.*

Proposición 5.1.10. *Sea H un espacio de Hilbert y M una variedad lineal de H , entonces M^\perp es un subespacio cerrado de H .*

Teorema 5.1.11 (Teorema de Hahn-Banach). *Sea E un espacio vectorial real y M un subespacio de E . Supongamos que p es un funcional sublineal sobre E , y $f : M \rightarrow \mathbb{R}$ es una aplicación lineal tal que $f(x) \leq p(x)$ para todo $x \in M$. Entonces, existe una aplicación lineal $g : E \rightarrow \mathbb{R}$ tal que $g|_M = f$ y $g(x) \leq p(x)$ para todo $x \in E$.*

Teorema 5.1.12 (Teorema de Banach-Steinhaus). *Supongamos que X e Y son espacios normados y que $A \subset L(X, Y)$. Consideremos las siguientes propiedades:*

(a) *La familia A es puntualmente acotada, es decir,*

$$\sup_{\Lambda \in A} \|\Lambda(x)\| < \infty \quad \text{para cada } x \in X.$$

(b) *El conjunto $B = \{x \in X : \sup_{\Lambda \in A} \|\Lambda(x)\| < \infty\}$ es de segunda categoría.*

(c) *La familia A está uniformemente acotada, esto es, $\sup_{\Lambda \in A} \|\Lambda(x)\| < \infty$. Entonces (b) implica (c), y (c) implica (a). Si, además, X es un espacio de Banach, entonces las tres propiedades (a), (b) y (c) son equivalentes.*

Teorema 5.1.13 (Teorema de Liouville). *Sea $f : \mathbb{C} \rightarrow \mathbb{C}$ una función entera y acotada, es decir, existe $M > 0$ tal que*

$$|f(z)| \leq M \quad \forall z \in \mathbb{C}$$

entonces resulta que f es constante.

Teorema 5.1.14 (Caracterización de mejor aproximación de conjuntos convexos). *Sea K un subconjunto convexo de un espacio escalar X , $x \in X$ e $y_0 \in K$. Entonces $y_0 = P_K(x)$ si y solo si*

$$\langle x - y_0, y - y_0 \rangle \leq 0 \quad \forall y \in K.$$

5.2. Teoría elemental de las álgebras de Banach

Para esta parte del apéndice hemos utilizado el libro [7].

Definición 5.2.1. Un álgebra compleja es un espacio vectorial A sobre los complejos en el que se define una multiplicación asociativa y distributiva,

$$x(yz) = (xy)z, \quad (x + y)z = xz + yz, \quad x(y + z) = xy + xz$$

para $x, y, z \in A$ y que está relacionado con la multiplicación por un escalar

$$\alpha(xy) = x(\alpha y) = (\alpha x)y$$

para $x, y \in A$.

Decimos que A es un álgebra compleja normada si hay una norma definida en A y si satisface la siguiente desigualdad:

$$\|xy\| \leq \|x\| \|y\| \quad (5.1)$$

$\forall x, y \in A$.

La desigualdad anterior convierte a la multiplicación en una operación continua, lo que significa que si $x_n \rightarrow x$ e $y_n \rightarrow y$, $x_n y_n \rightarrow xy$ y tenemos la siguiente igualdad

$$x_n y_n - xy = (x_n - x)y_n + x(y_n - y).$$

Además A tiene una unidad, es decir, que existe un elemento e tal que

$$xe = ex = x \quad (x \in A).$$

Dicho elemento unidad es único ($e' = e'e = e$) y $\|e\| \geq 1$.

Además A tiene un único elemento invertible x si existe un elemento x^{-1} tal que

$$x^{-1}x = xx^{-1} = e.$$

Definición 5.2.2. El espectro de un elemento $x \in A$, $\sigma(x)$, es el conjunto de todos los números complejos λ tal que $x - \lambda e$ es no invertible.

Teorema 5.2.3. Sea $x \in A$ y $\|x\| < 1$, entonces $e + x \in G$ (conjunto de todos los elementos invertibles de A)

$$(e + x)^{-1} = \sum_{n=0}^{\infty} (-1)^n x^n \quad (5.2)$$

$$\|(e + x)^{-1} - e + x\| \leq \frac{\|x\|^2}{1 - \|x\|}. \quad (5.3)$$

Demostración. La desigualdad se satisface por que la norma cumple que $\|x^n\| \leq \|x\|^n$. Si

$$s_N = e - x + x^2 - \cdots + (-1)^N x^N$$

$\{x_n\}$ es una sucesión de Cauchy en A . Como la multiplicación es continua y

$$(e + x)s_N = e + (-1)^N x^{N+1} = s_N(e + x)$$

observamos que $(e + x)y = e = y(e + x)$. Lo que da (5.2). Para obtener (5.3) tenemos

$$\left\| \sum_{n=2}^{\infty} (-1)^n x^n \right\| \leq \sum_{n=2}^{\infty} \|x^n\| \leq \sum_{n=2}^{\infty} \|x\|^n = \frac{\|x\|^2}{1 - \|x\|}.$$

□

Teorema 5.2.4. Si $x \in G$, $\|x^{-1}y\| = 1/\alpha$, $h \in A$, y $\|h\| < \beta < \alpha$. Luego $x + h \in G$ y

$$\|(x + h)^{-1} - x^{-1} + x^{-1}hx^{-1}\| \leq \frac{\beta^2}{\alpha^2(\alpha - \beta)}. \quad (5.4)$$

Demostración. $\|x^{-1}h\| \leq \beta/\alpha < 1$, por lo tanto, $e + x^{-1}h \in G$, por el teorema 5.2.3 y ya que $x + h = x(e + x^{-1}h)^{-1}x^{-1}$, tenemos que $x + h \in G$ y

$$(x + h)^{-1} = (e + x^{-1}h)^{-1}x^{-1}.$$

Así

$$(x + h)^{-1} - x^{-1} + x^{-1}hx^{-1} = [(e + x^{-1}h)^{-1} - e + x^{-1}h]x^{-1}.$$

La desigualdad (5.4) proviene del teorema 5.2.3 sustituyendo x por $x^{-1}h$. □

Corolario 5.2.5. G es un conjunto abierto y la aplicación $x \rightarrow x^{-1}$ es un homomorfismo que va de G a G .

Demostración. Para ello, sea $x \in G$ y $\|h\| \rightarrow 0$, (5.4) implica que

$$\|(x + h)^{-1} - x^{-1}\| \rightarrow 0.$$

Por lo tanto, $x \rightarrow x^{-1}$ es continua, es una aplicación que va de G a G y es su propia inversa, por lo que tenemos un homomorfismo. □

Corolario 5.2.6. La aplicación $x \rightarrow x^{-1}$ es diferenciable. Su diferencial para cualquier $x \in G$ es el operador lineal que lleva $h \in A$ a $-x^{-1}hx^{-1}$.

Corolario 5.2.7. Para cada $x \in A$, $\sigma(x)$ es compacto y $|\lambda| \leq \|x\|$ si $\lambda \in \sigma(x)$.

Demostración. Por reducción a lo absurdo, si $|\lambda| \geq \|x\|$, luego $e - \lambda^{-1}x \in G$, por el teorema 5.2.3, sabiendo que es cierto también para $x - \lambda e = -\lambda(e - \lambda^{-1}x)$; entonces $\lambda \notin \sigma(x)$. Para probar que $\sigma(x)$ es cerrado observemos que (a) $\lambda \in \sigma(x)$ si y solo si $x - \lambda e \notin G$, (b) el complementario de G es un subconjunto cerrado de A , por el corolario 5.2.5; y (c) la aplicación $\lambda \rightarrow x - \lambda e$ es una aplicación continua en el plano complejo en A . \square

Teorema 5.2.8. *Sea ϕ un funcional lineal y acotado en A , fijamos $x \in A$ y definimos*

$$f(\lambda) = \phi[(x - \lambda e)^{-1}] \quad (\lambda \notin \sigma(x)).$$

Entonces, f es holomorfa en el complementario de $\sigma(x)$ y $\lambda \rightarrow 0$ cuando $n \rightarrow \infty$.

Demostración. Fijamos $\lambda \notin \sigma(x)$ y aplicamos el teorema 5.2.4 con $x - \lambda e$ en lugar de x y con $(\lambda - \mu)e$ en lugar de h . Veamos que hay un constante C , que depende de λ y x , tal que

$$\|(x - \mu e)^{-1} - (x - \lambda e)^{-1} + (\lambda - \mu)(x - \lambda e)^{-2}\| \leq C|\mu - \lambda|^2$$

para todo μ que esté suficiente cerca de λ . Por lo tanto,

$$\frac{(x - \mu e)^{-1} - (x - \lambda e)^{-1}}{\mu - \lambda} \rightarrow (x - \lambda e)^{-2} \quad (5.5)$$

cuando $\mu \rightarrow \lambda$. Si aplicamos ϕ en ambos lados de (5.5), la continuidad y linealidad de ϕ muestra que

$$\frac{f(\mu) - f(\lambda)}{\mu - \lambda} \rightarrow \phi[(x - \lambda e)^{-2}].$$

Así que f es diferenciable y por lo tanto, holomorfa fuera de $\sigma(x)$. Finalmente, cuando $\lambda \rightarrow \infty$, obtenemos

$$\lambda f(\lambda) = \phi[\lambda(x - \lambda e)^{-1}] = \phi[(x/\lambda - e)^{-1}] \rightarrow \phi(-e)$$

por la continuidad de la aplicación inversa en G . \square

Teorema 5.2.9. *Para todo $x \in A$, $\sigma(x)$ es compacto y no vacío.*

Demostración. Ya sabemos que $\sigma(x)$ es compacto. Fijamos $x \in A$ y $\lambda_0 \notin \sigma(x)$. Entonces, por $(x - \lambda_0 e)^{-1} \neq 0$ y el teorema de Hahn-Banach, obtenemos la existencia de un funcional ϕ lineal y acotado en A tal que $f(\lambda_0) \neq 0$, con f definido como en el teorema 5.2.8. Si $\sigma(x)$ fuese vacío, el teorema 5.2.8 implica que f es una función completa que tiende de 0 a ∞ , por tanto $f(\lambda) = 0$ para todo λ , por el teorema de Liouville, y esto contradice que $f(\lambda_0) \neq 0$. Así concluimos con que $\sigma(x)$ no vacío. \square

Teorema 5.2.10 (Gelfand-Mazur). *Si A es un álgebra de Banach compleja con una unidad en la que cada elemento distinto de cero es invertible, entonces, A es el campo complejo. Un álgebra en la que cada elemento invertible es distinto de cero, se llama álgebra de división.*

Demostración. Si $x \in A$ y $\lambda_1 \neq \lambda_2$, al menos uno de los elementos $x - \lambda_1 e$ y $x - \lambda_2 e$ deben ser invertibles, ya que ambos no pueden ser cero. Se sigue del teorema 5.2.9 que $\sigma(x)$ esta formado por un solo punto, $\lambda(x)$, para cada $x \in A$. Como $x - \lambda(x)e$ es no invertible, debe ser 0, por tanto, $x = \lambda(x)e$. La aplicación $x \rightarrow \lambda(x)$ es aquí un isomorfismo de A en el plano complejo, que además es una isometría, ya que $|\lambda(x)| = \|\lambda(x)e\| = \|x\|$ para todo $x \in A$. \square

Definición 5.2.11. Para cualquier $x \in A$, el radio espectral $\rho(x)$ de x es el radio del disco cerrado más pequeño con centro en el origen que contiene a $\sigma(x)$:

$$\rho(x) = \sup\{|\lambda| : \lambda \in \sigma(x)\}$$

Teorema 5.2.12 (Fórmula del radio espectral). *Para cada $x \in A$,*

$$\lim_{n \rightarrow \infty} \|x^n\|^{1/n} = \rho(x)$$

Demostración. Fijamos $x \in A$, sea n un entero positivo, λ un número complejo y asumimos que $\lambda^n \notin \sigma(x^n)$. Tenemos

$$(x^n - \lambda^n e) = (x - \lambda e)(x^{n-1} + \lambda x^{n-2} + \dots + \lambda^{n-1} e).$$

Multiplicando ambos lados por $(x^n - \lambda^n e)^{-1}$, demostramos que $x - \lambda e$ es invertible, por tanto $\lambda \notin \sigma(x)$.

Luego si $\lambda \in \sigma(x)$, $\lambda^n \in \sigma(x^n)$ para $n = 1, 2, \dots$. El corolario 5.2.7 y el teorema 5.2.4 prueban que $|\lambda^n| \leq \|x^n\|$ y por lo tanto, $|\lambda| \leq \|x^n\|^{1/n}$.

Esto da

$$\rho(x) \leq \liminf_{n \rightarrow \infty} \|x^n\|^{1/n}. \quad (5.6)$$

Ahora, si $|\lambda| > \|x\|$ verificamos que

$$(\lambda e - x) \sum_{n=0}^{\infty} \lambda^{-n-1} x^n = e. \quad (5.7)$$

La serie anterior da $-(x - \lambda e)^{-1}$. Sea ϕ un funcional lineal y acotado de A y definimos f como en el teorema 5.2.8. Por (5.7), la expansión

$$f(\lambda) = - \sum_{n=0}^{\infty} \phi(x^n) \lambda^{-n-1} \quad (5.8)$$

es cierta para todo λ tal que $|\lambda| > \|x\|$. Por el teorema 5.2.8, f es holomorfa fuera de $\sigma(x)$, por lo tanto también en el conjunto $\{\lambda : |\lambda| > \rho(x)\}$. De aquí obtenemos que la serie de (5.8) converge si $|\lambda| > \rho(x)$. En particular,

$$\sup_n |\phi(\lambda^{-n} x^n)| < \infty \quad (|\lambda| > \rho(x)) \quad (5.9)$$

para cada funcional ϕ lineal y acotado de A .

Esto es consecuencia del teorema de Hahn-Banach, que la norma de cualquier elemento de A es la misma que su norma como un funcional lineal en el espacio dual de A . Ya que (5.9) se cumple para todo ϕ , podemos aplicar el teorema de Banach-Steinhaus y concluir que para cada λ con $|\lambda| > \rho(x)$ se corresponde un número real $C(\lambda)$ tal que

$$\|\lambda^{-n}x^n\| \leq C(\lambda) \quad (n = 1, 2, \dots). \quad (5.10)$$

Multiplicando (5.10) por $|\lambda|^n$ y tomando la raíz n -ésima, tenemos

$$\|x^n\|^{1/n} \leq |\lambda|[C(\lambda)]^{1/n} \quad (n = 1, 2, \dots). \quad (5.11)$$

si $|\lambda| > \rho(x)$ y por tanto,

$$\limsup_{n \rightarrow \infty} \|x^n\|^{1/n} \leq \rho(x). \quad (5.12)$$

El teorema concluye con (5.6) y (5.11). \square

5.3. El teorema espectral para operadores hermíticos

En esta sección nos vamos a basar en el artículo [5].

Sea ϕ una función medible, acotada y compleja en un espacio medible X con medida μ . Se define el operador A en $L^2(\mu)$ como $(Af)(x) = \phi(x)f(x)$, este operador es la multiplicación inducida por ϕ . Si X es finito entonces A es diagonal.

Notación 5.3.1. $A\langle f_1, \dots, f_n \rangle = \langle \phi_1 f_1, \dots, \phi_n f_n \rangle$.

Teorema 5.3.2 (Teorema espectral). *Todo operador hermítico es unitariamente equivalente a una multiplicación. Es decir, si A hermítico en un espacio de Hilbert H , entonces existe una función ϕ real, medible, acotada en un espacio medible con medida μ y existe una isometría*

$$U : L^2(\mu) \rightarrow H \text{ tal que} \\ (U^{-1}AUf)(x) = \phi(x)f(x) \quad x \in X, f \in L^2(\mu).$$

Herramientas para la demostración:

1) Definición de espectro

Definición 5.3.3. Sea z un número complejo, se dice que dicho z pertenece al espectro de un operador A si el operador $A - z$ no es invertible.

En dimensión finita, esto equivale a decir que $A - z$ no es inyectivo de modo que z es un valor propio, por tanto, z pertenece al espectro puntual. Pero en dimensión infinita, $A - z$ puede ser inyectivo pero no sobreyectivo. En general, hay más elementos en el espectro de un operador que sus valores propios.

2) Igualdad de la norma y radio espectral

Definición 5.3.4. Si el espectro de A es $\Lambda(A)$, el radio de A se define como

$$\rho(A) = \sup\{|\lambda| : \lambda \in \Lambda(A)\}.$$

Definamos ahora los conceptos de valor propio, valor propio aproximado y espectro puntual aproximado que nos servirán para explicar la próxima proposición.

Definición 5.3.5. λ es un valor propio de A si existe un vector unitario tal que

$$\|Ax - \lambda x\| = 0.$$

Definición 5.3.6. Diremos que un número complejo λ es un valor propio aproximado de A si para cada número positivo ϵ existe un vector unitario x tal que $\|Ax - \lambda x\| < \epsilon$. Esto es equivalente a decir que $\forall \epsilon$ existe un vector $x \neq 0$ tal que $\|Ax - \lambda x\| < \epsilon \|x\|$.

Definición 5.3.7. El espectro puntual aproximado de un operador A , $\pi(A)$, es el conjunto de valores propios aproximados de A .

Para la demostración de la siguiente proposición nos hemos basado en el libro [6].

Proposición 5.3.8.

- a) Siempre es cierto que $\rho(A) \leq \|A\|$ para cualquier operador A (sea normal o no).
 b) Si además usamos que A es hermítico, tenemos que $\rho(A) = \|A\|$.

Demostración.

- a) Si $\lambda_0 \in \Lambda(A)$, entonces $A - \lambda_0$ invertible, luego

$$\|1 - (A - \lambda_0)^{-1}(A - \lambda)\| = \|(A - \lambda_0)^{-1}((A - \lambda_0)(A - \lambda))\| \leq \|(A - \lambda_0)^{-1}\| \cdot \|\lambda - \lambda_0\|$$

En consecuencia, $\|1 - (A - \lambda_0)^{-1}(A - \lambda)\| \leq 1$ para $|\lambda - \lambda_0|$ suficientemente pequeño. Como $(A - \lambda_0)^{-1}(A - \lambda)$ es invertible siempre que $|\lambda - \lambda_0|$ sea suficientemente pequeño. Esto significa que el complementario de $\Lambda(A)$ es un subconjunto abierto del plano complejo. Si $|\lambda| > \|A\|$, entonces $\|A/\lambda\| < 1$, por lo que $1 - (A/\lambda)$ es invertible. De aquí se deduce que $\lambda \in \Lambda(A)$ y que si $\lambda \in \Lambda(A)$, entonces $|\lambda| \leq \|A\|$.

- b) Para probar la igualdad debemos demostrar que $\|A\|^2 \in \pi(A^2)$, podemos concluir (por el teorema de la aplicación espectral demostrado próximamente) que $\pm\|A\| \in \Lambda(A)$. La prueba se basa en la identidad $\|A^2x - \lambda^2x\|^2 = \|A^2x\|^2 - 2\lambda^2\|Ax\|^2 + \lambda^4\|x\|^2$ válida (ya que A es hermítico) $\forall \lambda$ real y $\forall x$ vector. Si $\{x_n\}$ es una sucesión de vectores unitarios tal que $\|Ax_n\| \rightarrow \|A\|$ y si $\lambda = \|A\|$ tenemos que

$$\|A^2x_n - \lambda^2x_n\|^2 \leq (\|A\| \cdot \|Ax_n\|)^2 - 2\lambda^2\|Ax_n\|^2 + \lambda^4 = \lambda^4 - \lambda^2\|Ax_n\|^2 \rightarrow 0.$$

Por lo que tenemos que $\|A\|^2 \in \pi(A^2)$.

□

3) Teorema de representación de Riesz para funcionales lineales

Teorema 5.3.9. *Sea L un funcional lineal positivo para todas las funciones continuas de valor real en un subconjunto compacto lineal X , entonces existe una única medida finita μ en los conjuntos Borel de X , tal que $L(f) = \int f d\mu \quad \forall f$ en el dominio de L .*

4) Teorema de aproximación de Weierstrass para funciones continuas

Teorema 5.3.10. *Cada función continua de valor real en un subconjunto compacto lineal real es el límite uniforme de los polinomios.*

5) Teorema de la aplicación espectral

Teorema 5.3.11. *Sea A un operador y p un polinomio, tenemos que*

$$\Lambda(p(A)) = p(\Lambda(A)) = \{p(\lambda) : \lambda \in \Lambda(A)\}.$$

Demostración.

\supseteq . Para cualquier número complejo λ_0 , existe un polinomio q tal que

$$p(\lambda) - p(\lambda_0) = (\lambda - \lambda_0)q(\lambda).$$

Entonces $p(A) - p(\lambda_0) = (A - \lambda_0)q(A)$ es no invertible. Esto significa que $p(A) - p(\lambda_0)$ es no invertible y $p(\lambda_0) \in \Lambda(p(A))$, por lo que $p(\Lambda(A)) \subseteq \Lambda(p(A))$.

\subseteq . Supongamos ahora que $\lambda_0 \in \Lambda(p(A))$ y que $\lambda_1, \dots, \lambda_n$ son raíces de la ecuación $p(\lambda) = \lambda_0$. Por tanto, $p(A) - \lambda_0 = \alpha(A - \lambda_1) \cdots (A - \lambda_n)$ para algún número complejo $\alpha \neq 0$, entonces $A - \lambda_j$ no puede ser invertible para al menos un valor de j ($1 \leq j \leq n$). Para dicho valor, tenemos que $\lambda_j \in \Lambda(A)$ y $p(\lambda_j) = \lambda_0$, así que $\lambda_0 \in p(\Lambda(A))$ y por lo tanto, $\Lambda(p(A)) \subseteq p(\Lambda(A))$. \square

Tras esto, ya podemos empezar con la demostración del teorema.

Demostración. Consideremos ahora un operador hermítico A en un espacio de Hilbert H . Un vector ξ en H es un vector cíclico para A si el conjunto de todos los vectores de la forma $q(A)\xi$, donde q recorre sobre los polinomios con coeficientes complejos, es denso en H .

Los vectores cíclicos pueden no existir, pero un argumento transfinito muestra que H es siempre la suma directa de una familia de subespacios, cada uno de los cuales reduce A de modo que la restricción de A a cada uno de ellos tiene un vector cíclico.

Una vez conocido el teorema espectral para cada restricción, se sigue para el propio A ; los espacios de medida que sirven para los sumandos directos de H tienen una suma directa natural, que valen para el propio H . En conclusión, no hay pérdida de generalidad al

suponer que A tiene un vector cíclico, ξ .
Para cada polinomio real p , escribimos

$$L(p) = \langle p(A)\xi, \xi \rangle$$

siendo L un funcional lineal. Entonces

$$\begin{aligned} |L(p)| &\leq \|p(A)\| \cdot \|\xi\|^2 = \rho(p(A)) \cdot \|\xi\|^2 \\ &= \sup\{|\lambda| : \lambda \in \Lambda(p(A))\} \cdot \|\xi\|^2 \\ &= \sup\{|p(\lambda)| : \lambda \in \Lambda(A)\} \cdot \|\xi\|^2 \end{aligned}$$

el funcional L está acotado por polinomios. Se deduce (por el teorema de Weierstrass) que L tiene una extensión acotada a todas las funciones continuas de valor real en $\Lambda(A)$. Para probar que L es positivo, veamos primero que si p es un polinomio real, entonces

$$\langle p(A)^2\xi, \xi \rangle = \|p(A)\xi\|^2 \geq 0.$$

Si f es una función arbitraria continua y positiva de $\Lambda(A)$, aproximamos \sqrt{f} uniformemente por polinomios reales, esta inecuación probada implica que $L(f) \geq 0$ (ya que f se aproxima uniformemente por cuadrados de polinomios reales).

El teorema de Riesz implica la existencia de una medida finita μ tal que

$$\langle p(A)\xi, \xi \rangle = \int p \, d\mu$$

para cada polinomio real p .

Para cada polinomio q (posiblemente complejo), definimos

$$Uq = q(A)\xi.$$

Como A es hermítico, $(q(A))^* = \bar{q}(A)$ es un polinomio de A y también lo es $(q(A))^*q(A) = |q|^2(A)$, de lo que se deduce que

$$\int |q|^2 d\mu = \langle \bar{q}(A)q(A)\xi, \xi \rangle = \langle (q(A))^*q(A)\xi, \xi \rangle = \|q(A)\xi\|^2 = \|Uq\|^2.$$

Esto significa que la transformación lineal U de un subconjunto denso de $L^2(\mu)$ en H es una isometría y por lo tanto, tiene una única extensión isométrica que va de $L^2(\mu)$ en H . La suposición de que ξ es un vector cíclico implica que el rango de U es denso e igual a H . Queda demostrar que la transformación de A por U es una multiplicación. Sea $\phi(\lambda) = \lambda \, \forall \lambda \in \Lambda(A)$ y dado un polinomio complejo q , tenemos que

$$\tilde{q}(\lambda) = \lambda q(\lambda) = \phi(\lambda)q(\lambda).$$

Entonces:

$$U^{-1}AUq = U^{-1}Aq(A)\xi = U^{-1}\tilde{q}(A)\xi = U^{-1}U\tilde{q} = \tilde{q}.$$

Concluimos con que $U^{-1}AU$ es igual a la multiplicación inducida por ϕ . Esto completa la prueba del teorema espectral para operadores hermíticos. \square

Corolario 5.3.12. *Asumiendo las hipótesis del teorema espectral, si T es además positivo, entonces la función F del teorema espectral es no negativo (c.t.p μ).*

Demostración. Sea $f \in L_2(\Omega, \mu)$ arbitrario e $y = U^{-1}f$. Como T es positivo, tenemos:

$$\int_{\Omega} F|f|^2 d\mu = \langle Ff, f \rangle = \langle Df, f \rangle = \langle UTU^{-1}f, f \rangle = \langle TU^{-1}f, U^*f \rangle = \langle Ty, y \rangle \geq 0$$

Por lo que $\int_{\Omega} F|f|^2 d\mu \geq 0$ para todo $f \in L_2(\Omega, \mu)$. Deducimos que $F \geq 0$ (c.t.p μ). \square

Observación 5.3.13. Esta demostración del teorema espectral nos sirve porque lo hemos aplicado a un operador positivo, que es un caso especial de operador hermítico.

Bibliografía

- [1] Heinz H. Bauschke, Frank Deutsch, and Hein Hundal. Characterizing arbitrarily slow convergence in the method of alternating projections. *Int. Trans. Oper. Res.*, 16(4):413–425, 2009.
- [2] Frank Deutsch. The method of alternating orthogonal projections. In *Approximation theory, spline functions and applications (Maratea, 1991)*, volume 356 of *NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci.*, pages 105–121. Kluwer Acad. Publ., Dordrecht, 1992.
- [3] Frank Deutsch. *Best approximation in inner product spaces*, volume 7 of *CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC*. Springer-Verlag, New York, 2001.
- [4] René Escalante. Algoritmos numéricos para el problema de restauración de imágenes usando el método de las proyecciones alternantes. *Revista de Matemática: Teoría y Aplicaciones*, 11(1):41–58, 2004.
- [5] P. R. Halmos. What does the spectral theorem say? *Amer. Math. Monthly*, 70:241–247, 1963.
- [6] Paul R. Halmos. *Introduction to Hilbert Space and the theory of Spectral Multiplicity*. Chelsea Publishing Company, New York, N. Y., 1951.
- [7] Walter Rudin. *Real and complex analysis*. McGraw-Hill Book Co., New York, third edition, 1987.