



Universidad de Sevilla

Facultad de Matemáticas

Trabajo Fin de Estudios

Grado en Matemáticas

Sobre problemas variacionales y optimización
(On variational problems and optimization)

Laura María Rodríguez Llave

Tutor

Faustino Maestre Caballero

Dpto. Ecuaciones Diferenciales y Análisis Numérico.

Abstract

The Calculus of Variations is a mathematical field that tries to find the functions that minimize or maximize the value of certain functional defined on a determined space. It is a classic topic in mathematics, interconnected with many other problems in this branch. It has multiple applications, among them is the use in optimal control.

In this work we will analyze optimization problems within the framework of the Calculus of Variations, as well as the resolution of some partial differential equations.

We will study the theory of existence and uniqueness of solution of certain optimization problems and we will see necessary and/or sufficient conditions of optimality, which will be important both from the theoretical and numerical point of view and they will allow us to characterize the optimum. In addition, we will describe numerical methods to solve minimization problems without constraints.

Resumen

El Cálculo de Variaciones es un campo de las matemáticas que trata de encontrar las funciones que minimizan o maximizan el valor de un cierto funcional definido sobre un determinado espacio. Es un tema clásico en matemáticas, interconectado con muchos otros problemas de esta rama. Tiene múltiples aplicaciones, entre ellas está el uso en control óptimo.

En este trabajo analizaremos problemas de optimización dentro del marco del Cálculo de Variaciones, así como la resolución de algunas ecuaciones en derivadas parciales no lineales.

Estudiaremos la teoría de existencia y unicidad de solución de ciertos problemas de optimización y veremos condiciones necesarias y/o suficientes de optimalidad, que serán importantes tanto desde el punto de vista teórico como numérico y nos permitirán caracterizar los óptimos.

Además, describiremos métodos numéricos para la resolución de problemas de minimización sin restricciones.

Índice

Introducción	3
1. Resultados previos de Análisis Funcional	5
1.1. Teoremas de Representación Riesz, de Lax-Milgram y de la Proyección	5
1.2. Resultados auxiliares sobre espacios de funciones	8
1.3. Distribuciones	10
1.4. Los espacios de Sobolev : $H^1(\Omega)$, $H_0^1(\Omega)$ y $H^{-1}(\Omega)$	14
1.5. Trazas y fórmulas de Green	17
1.6. Resultado de compacidad	18
1.7. Otros espacios : $H^m(\Omega)$, $H(\text{div})$ y $W^{m,p}(\Omega)$	19
1.8. Convergencia débil	21
2. Optimización	23
2.1. Introducción	23
2.2. Ejemplos	24
2.2.1. En dimensión finita	24
2.2.2. En dimensión infinita	26
2.3. Convexidad	27
3. Optimización diferenciable	34
3.1. La diferencial de Fréchet y de Gâteaux	34
3.2. Condiciones de optimalidad (caso convexo)	39
4. Métodos numéricos en optimización sin restricciones	42
4.1. Algoritmo de gradiente con paso óptimo	42
4.2. Algoritmo de gradiente con paso fijo	45
4.3. Método del gradiente conjugado para funcionales cuadráticos	46
5. Aplicación: resolución edp no lineal	51
5.1. Marco teórico para la solución del problema	51
5.2. Resolución del problema con FreeFEM	57
5.2.1. En $\Omega = B(0, 1)$	57
5.2.2. En Ω una elipse con un agujero	62
Anexo	66
Bibliografía	74

Índice de figuras

2.1. Fases del proceso de optimización	24
2.2. Conjunto no convexo	30
2.3. Conjunto convexo	30
5.1. Frontera del dominio	59
5.2. Malla del dominio	59
5.3. Función f	59
5.4. Función u_0 : inicialización AGPO	60
5.5. Primera iteracion del AGPO	60
5.6. Segunda iteracion del AGPO	60
5.7. Tercera iteracion del AGPO	61
5.8. 4 ^a iteración del Algoritmo de gradiente con paso fijo	61
5.9. 17 ^a iteración del Algoritmo de gradiente con paso fijo	62
5.10. 38 ^a iteración del Algoritmo de gradiente con paso fijo	62
5.11. Frontera elipse	63
5.12. Malla de la elipse	63
5.13. Función f_1	64
5.14. Función u_0 : inicialización AGPO (para la elipse y f_1)	64
5.15. Primera iteracion del AGPO (para la elipse y f_1)	64
5.16. Segunda iteracion del AGPO (para la elipse y f_1)	65
5.17. Tercera iteracion del AGPO (para la elipse y f_1)	65

Introducción

La optimización dentro del Cálculo de Variaciones se enfoca en encontrar la función que optimiza un cierto funcional, sujeta a ciertas restricciones o condiciones. Este área de investigación tiene aplicaciones en diversos campos, como la física, la economía, la ingeniería y la biología, entre otros. Al encontrar la función que minimiza o maximiza un funcional, se pueden obtener soluciones óptimas para problemas de diseño, control, estimación o predicción.

La resolución de ecuaciones en derivadas parciales y la optimización en el Cálculo de Variaciones a menudo se entrelazan, ya que muchos problemas prácticos requieren encontrar funciones que optimicen ciertos criterios mientras satisfacen ciertas ecuaciones diferenciales. Estos problemas son abordados utilizando métodos numéricos y técnicas de optimización, como el método de los elementos finitos, el método de los mínimos cuadrados, los algoritmos de gradiente descendente y los métodos de optimización convexa.

La presente Memoria está organizada como sigue:

El Capítulo 1, recoge los conceptos básicos y algunos resultados fundamentales de Análisis Funcional que serán necesarios para desarrollar esta memoria.

Entre ellos, los teoremas clásicos de Representación de Riesz y de Lax-Milgram. Introduciremos la teoría de distribuciones, discutiremos ciertos aspectos esenciales sobre la teoría de Espacios de Sobolev y trataremos la compacidad y la convergencia débil. Todo ello nos permitirá abordar diversos problemas variacionales.

En el Capítulo 2, se presentarán algunos ejemplos de optimización que provienen de diversos ámbitos. Además, estudiaremos el concepto de funcional elíptico (α -convexo) que será necesario en el siguiente capítulo. Trataremos principalmente la cuestión de existencia y unicidad en optimización continua, donde será fundamental el papel de la convexidad y distinguiremos entre problemas de optimización en dimensión finita e infinita, ya que requieren ser tratados de distinta forma.

El Capítulo 3, está dedicado a la teoría de la diferenciabilidad en dimensiones infinitas, centrándonos en las diferenciales de Fréchet y Gâteaux, tanto de primer como de segundo orden.

Discutiremos sus definiciones, propiedades y relaciones, así como su aplicación en la resolución de problemas prácticos. Finalmente, estudiaremos condiciones de optimalidad en el caso en el que el conjunto de restricciones sea convexo, las cuales juegan un papel fundamental al determinar las soluciones óptimas.

En el Capítulo 4, se estudian varios de los métodos más importantes para la resolución de problemas de optimización sin restricciones. Estos métodos combinan principios matemáticos, algoritmos iterativos y capacidades computacionales para aproximarse a la solución óptima de manera eficiente y precisa. En particular, dos métodos numéricos destacados en la optimización sin restricciones son el método de gradiente y el método de direcciones conjugadas.

En el Capítulo 5, trabajaremos con un funcional elíptico concreto, consideraremos el problema de mínimos correspondiente y lo resolveremos mediante uno de los métodos numéricos estudiados en el capítulo anterior, comprobando previamente que se cumplen las hipótesis necesarias. Para ello, haremos un programa en FreeFEM.

Capítulo 1

Resultados previos de Análisis Funcional

1.1. Teoremas de Representación Riesz, de Lax-Milgram y de la Proyección

Recordemos primero la definición de espacio de Banach y espacio de Hilbert.

Definición 1.1 Diremos que V es un espacio de Banach si es un espacio vectorial normado y completo.

Definición 1.2 Diremos que H es un espacio de Hilbert si es un espacio vectorial con un producto escalar cuya norma asociada es completa.

Notación:

- X, Y, \dots denotan espacios normados sobre \mathbb{R} (espacio vectorial sobre \mathbb{R} con una norma asociada).
- Denotaremos $\|\cdot\|_X, \|\cdot\|_Y, \dots$ las respectivas normas (o simplemente $\|\cdot\|$).
- H, V, \dots denotan espacios pre-hilbertianos sobre \mathbb{R} (espacio vectorial sobre \mathbb{R} con un producto escalar).
- Denotaremos $\langle \cdot, \cdot \rangle_H, \langle \cdot, \cdot \rangle_V, \dots$ los respectivos productos escalares (o simplemente $\langle \cdot, \cdot \rangle$).

Teorema de Riesz

Sean X, Y dos espacios normados sobre \mathbb{R} . Consideraremos aplicaciones $T : X \longrightarrow Y$ lineales y continuas, lo que llamaremos como operador lineal y continuo.

Definición 1.3 Diremos que la aplicación T es lineal si

$$T(\alpha x + \beta y) = \alpha T(x) + \beta T(y), \quad \forall \alpha, \beta \in \mathbb{R}, \quad \forall x, y \in X.$$

Y será continua si

$$\forall \{x_n\}_{n \geq 1} \subset X \text{ tal que } x_n \longrightarrow x \text{ en } X \text{ se tiene } T(x_n) \longrightarrow T(x).$$

Definición 1.4 $\mathcal{L}(X, Y)$ es el espacio vectorial de los operadores lineales y continuos de X en Y . La norma del operador T será:

$$\|T\|_{\mathcal{L}(X, Y)} = \sup_{v \in X, v \neq 0} \frac{\|Tv\|_Y}{\|v\|_X} = \sup_{\|v\|_X \leq 1} \|Tv\|_Y = \sup_{\|v\|_X = 1} \|Tv\|_Y \quad (1.1)$$

Observación 1.5

1. Si $X = Y$, escribiremos $\mathcal{L}(X)$.
2. En el caso $Y = \mathbb{R}$, $\mathcal{L}(X, \mathbb{R}) \equiv X'$ espacio dual topológico de X . Siempre es un espacio de Banach.
3. De (1.1) deducimos $\|Tv\|_Y \leq \|T\|_{\mathcal{L}(X, Y)} \|v\|_X$, $\forall T \in \mathcal{L}(X, Y)$, $\forall v \in X$.

Teorema 1.6 (De representación de Riesz) Sea H un espacio de Hilbert con producto escalar $\langle \cdot, \cdot \rangle_H$ y norma $\|\cdot\|_H$. Sea H' su dual. Para toda forma lineal continua $L \in H'$ existe un único $y \in H$ tal que

$$L(x) = \langle y, x \rangle \quad \forall x \in H.$$

Además, se tiene $\|L\|_{H'} = \|y\|_H$.

Del Teorema de representación de Riesz se deduce el siguiente corolario importante.

Corolario 1.7 Sea H un espacio de Hilbert, $K \subset H$ un convexo cerrado no vacío y $f \in H'$. Entonces, $\exists!$ $u \in K$ solución del problema de mínimos

$$\begin{cases} \text{Minimizar} & \frac{1}{2}\|v\|_H^2 - f(v) \\ \text{sujeto a} & v \in K \end{cases} \quad (1.2)$$

Además, u está caracterizada por

$$\begin{cases} u \in K \\ \langle u, v - u \rangle_H \geq f(v - u) \quad \forall v \in K \end{cases} \quad (1.3)$$

Todos los resultados anteriores se detallan en [1].

Teorema de Lax-Milgram

Para obtener la existencia y unicidad de solución de una formulación variacional en un espacio de Hilbert es necesario dicho teorema. Podemos consultar los siguientes conceptos en [2].

Definición 1.8 Sea H un espacio de Hilbert y $a(\cdot, \cdot) : H \times H \longrightarrow \mathbb{R}$ una forma bilineal. Se dice que $a(\cdot, \cdot)$ es acotada si $\exists M > 0$ tal que $|a(u, v)| \leq M\|u\|\|v\|$, $\forall u, v \in H$.

Definición 1.9 Sea H un espacio de Hilbert y $a(\cdot, \cdot) : H \times H \rightarrow \mathbb{R}$ una forma bilineal. Diremos que $a(\cdot, \cdot)$ es coercitiva (o H -elíptica) si $\exists \alpha > 0$ tal que $a(v, v) \geq \alpha \|v\|^2$, $\forall v \in H$.

Proposición 1.10 La forma bilineal $a(\cdot, \cdot) : H \times H \rightarrow \mathbb{R}$ es acotada $\iff a(\cdot, \cdot)$ es continua en $H \times H$.

Proposición 1.11 Sean H Hilbert, $a : (u, v) \mapsto a(u, v) \in \mathbb{R}$ una forma bilineal y continua (con constante M), entonces $\exists!$ operador $A \in \mathcal{L}(H)$ tal que $\langle Au, v \rangle = a(u, v)$, $\forall u, v \in H$. Además $\|Au\| \leq M\|u\|$, $\forall u \in H$.

Teorema 1.12 (Lax-Milgram) Sean H un espacio de Hilbert, $a : H \times H \rightarrow \mathbb{R}$ una forma bilineal, continua y coercitiva, y sea $L : H \rightarrow \mathbb{R}$ una forma lineal, es decir,

$$\exists C > 0, \text{ tal que } |L(v)| \leq C\|v\|_H, \quad \forall v \in H.$$

Entonces existe una única solución del problema :

$$\begin{cases} \text{Hallar } u \in H \text{ tal que} \\ a(u, v) = L(v), \quad \forall v \in H. \end{cases} \quad (1.4)$$

Además, si $a(\cdot, \cdot)$ es simétrica, definimos $J(v) := \frac{1}{2}a(v, v) - L(v)$. Entonces la solución u de (1.4) se caracteriza por ser la única solución del problema de mínimos

$$\begin{cases} \text{Hallar } u \in H \text{ tal que} \\ J(u) \leq J(v), \quad \forall v \in H. \end{cases} \quad (1.5)$$

El Teorema de Lax-Milgram en particular dice que A es un isomorfismo.

Observación 1.13

La solución u depende continuamente del dato L : en efecto, si L_1 y L_2 son dos formas lineales continuas sobre H , es decir, dos elementos del dual H' y denotamos $\|\cdot\|_{H'}$ la norma en H' , entonces las soluciones asociadas u_1 y u_2 verifican:

$$\|u_1 - u_2\|_H \leq \frac{1}{\alpha} \|L_1 - L_2\|_{H'}. \quad (1.6)$$

Observación 1.14

El Teorema de Lax-Milgram nos da una condición suficiente para la existencia y unicidad de solución de (1.4). Pero no es una condición necesaria (ni siquiera en el caso de dimensión finita lo es).

Teorema de la Proyección

Para más detalles sobre la ortogonalidad, véase [3].

Definición 1.15 Sea H un espacio de Hilbert, $A \subset H$ un subconjunto no vacío. El ortogonal de A es el conjunto $A^\perp = \{v \in H : \langle u, v \rangle_H = 0, \forall u \in A\}$.

Proposición 1.16 Sea H un espacio de Hilbert. Se tiene:

1. $A, B \subset H, \quad A \subset B \implies B^\perp \subset A^\perp$.
2. $A \subset H \implies A^\perp \subset H$ es un subespacio vectorial cerrado de H .
3. $A^\perp = \overline{A}^\perp = \overline{[A]}^\perp$, ($[A]$ subespacio vectorial generado por los elementos de A).

Teorema 1.17 (De la Proyección) Sea H un espacio de Hilbert, $K \subset H$ convexo cerrado no vacío y $u_0 \in H$. Entonces, existe una única solución $\hat{u} \in K$ del problema de mejor aproximación

$$\begin{cases} \text{Hallar } \hat{u} \in K \\ \|\hat{u} - u_0\| = \inf_{v \in K} \|v - u_0\| \end{cases} \quad (1.7)$$

Además, \hat{u} está caracterizada por

$$\begin{cases} \hat{u} \in K \\ \langle u_0 - \hat{u}, v - \hat{u} \rangle \leq 0, \quad \forall v \in K \end{cases} \quad (1.8)$$

1.2. Resultados auxiliares sobre espacios de funciones

Para la elaboración de esta Sección, se han empleado principalmente las referencias [1] y [4].

Sea $\Omega \subset \mathbb{R}^N$ un abierto no vacío. Dada $\varphi \in C^0(\Omega)$, decimos que el soporte de φ es $\text{sop } \varphi = \overline{\{x \in \Omega : \varphi(x) \neq 0\}}$.

Por tanto deducimos que:

1. $\text{sop } \varphi \subset \overline{\Omega}$
2. Si $x \in \Omega \setminus \text{sop } \varphi \implies \varphi(x) = 0$

A partir de $\text{sop } \varphi$ definimos los siguientes conjuntos.

Definición 1.18 $C_0^0 = C_c^0 = \{\varphi \in C^0(\Omega) : \text{sop } \varphi \subset \Omega \text{ compacto}\}$ (funciones continuas y de soporte compacto).

$C_k^0(\Omega) = C^k(\Omega) \cap C_0^0(\Omega)$ (funciones de clase k y de soporte compacto).

$\mathcal{D}(\Omega) = C^\infty(\Omega) \cap C_0^0(\Omega)$ (funciones de clase ∞ y de soporte compacto).

$\mathcal{D}(\Omega)$ también se conoce como el espacio de funciones "test".

Definición 1.19 Consideramos un conjunto $\Omega \subset \mathbb{R}^N$ abierto, dotado de la medida de Lebesgue y sea $p \in [1, \infty]$. Definimos el espacio $L^p(\Omega)$ como:

$$L^p(\Omega) = \left\{ u : \Omega \longrightarrow \mathbb{R} \text{ medibles} : \int_{\Omega} |u(x)|^p dx < \infty \right\} \text{ si } p \in [1, \infty). \quad (1.9)$$

$$L^\infty(\Omega) = \{u : \Omega \longrightarrow \mathbb{R} \text{ medibles} : \exists C > 0 \text{ tal que } |u(x)| < C \text{ e.c.t } \Omega\}. \quad (1.10)$$

Observación 1.20

Los elementos de $L^p(\Omega)$ no son funciones sino clases de funciones: dos funciones iguales salvo en un conjunto de medida nula se consideran la misma. Esto significa que no podemos hablar del valor de una función de $L^p(\Omega)$ en un punto de Ω .

Definición 1.21 La norma asociada al espacio $L^p(\Omega)$ es :

1. Si $p \in [1, \infty)$,

$$\|u\|_{L^p(\Omega)} = \|u\|_{p;\Omega} = \left(\int_{\Omega} |u(x)|^p dx \right)^{1/p}, \quad \forall u \in L^p(\Omega). \quad (1.11)$$

2.

$$\|u\|_{L^\infty(\Omega)} = \|u\|_{\infty;\Omega} = \inf\{C > 0 : |u(x)| < C \text{ e.c.t. } \Omega\}, \quad \forall u \in L^\infty(\Omega). \quad (1.12)$$

Para el siguiente Teorema véase [5].

Teorema 1.22 $L^p(\Omega)$ es un espacio de Banach $\forall p \in [1, \infty]$. Si $p = 2$, $L^2(\Omega)$ es un espacio de Hilbert para el producto

$$\langle u, v \rangle_{L^2(\Omega)} = \int_{\Omega} u(x)v(x) dx, \quad \forall u, v \in L^2(\Omega) \quad (1.13)$$

Recordemos el siguiente resultado de densidad.

Teorema 1.23 Si $p \in [1, \infty)$, $\mathcal{D}(\Omega)$ es denso en $L^p(\Omega)$.

De dicho teorema obtenemos las siguientes consecuencias:

Observación 1.24

1. Si $u \in L^2(\Omega)$ es tal que $\int_{\Omega} u(x)\varphi(x) dx = 0 \quad \forall \varphi \in \mathcal{D}(\Omega) \implies u = 0$ e.c.t Ω .
2. Para cada $p \in [1, \infty)$, $L^p(\Omega)$ es un espacio de Banach separable (existe un subconjunto denso y numerable).

Definición 1.25 Sea v una función medible en $\Omega \subset \mathbb{R}^N$ abierto no vacío. Se dice que v es localmente integrable en Ω si $v|_K \in L^1(\Omega)$, $\forall K \subset \Omega$ compacto.

Así,

$$\begin{aligned} L^1_{Loc}(\Omega) &= \{v : v \text{ medible y localmente integrable en } \Omega\} \\ &= \{v : v \text{ medible tal que } \int_K |u(x)| dx < \infty, \forall K \text{ compacto en } \Omega\}. \end{aligned} \quad (1.14)$$

Proposición 1.26 Sea $\Omega \subset \mathbb{R}^N$ abierto no vacío. Entonces,

$$L^p(\Omega) \subset L^1_{Loc}(\Omega) \quad \forall p \in [1, \infty] \quad \text{y} \quad C^0(\Omega) \subset L^1_{Loc}(\Omega).$$

Teorema 1.27 (Fundamental del Cálculo Variacional) Sea $u \in L^1_{Loc}(\Omega)$, tal que

$$\int_{\Omega} u(x)\varphi(x) dx = 0, \quad \forall \varphi \in \mathcal{D}(\Omega). \quad \text{Entonces, } u(x) = 0 \quad \text{p.c.t. } x \in \Omega.$$

A continuación veamos algunas desigualdades que serán de gran utilidad. Para ello definimos el exponente conjugado.

Definición 1.28 Dado $p \in (1, +\infty)$, se llama exponente conjugado de p a $p' \in (1, +\infty)$ dado por

$$\frac{1}{p} + \frac{1}{p'} = 1 \tag{1.15}$$

Si $p = 1$, su exponente conjugado es $p' = \infty$. Del mismo modo, si $p = \infty$, su exponente conjugado es $p = 1$. En particular $p = 2 \iff p' = 2$.

Proposición 1.29 (Desigualdad de Young) Sea $p \in (1, +\infty)$ y p' su exponente conjugado. Entonces se cumple

$$ab \leq \frac{1}{p}a^p + \frac{1}{p'}b^{p'} \tag{1.16}$$

Proposición 1.30 (Desigualdad de Hölder) Sea $p \in [1, +\infty]$ y p' su exponente conjugado. Entonces

$$\int_{\Omega} |u(x)v(x)| dx \leq \|u\|_{p;\Omega} \|v\|_{p';\Omega} \quad \forall u \in L^p(\Omega), \forall v \in L^{p'}(\Omega) \tag{1.17}$$

Proposición 1.31 (Desigualdad de Minkowski) Sea $p \in [1, +\infty)$ y $u, v \in L^p(\Omega)$. Entonces

$$\left(\int_{\Omega} |u(x) + v(x)|^p dx \right)^{1/p} \leq \left(\int_{\Omega} |u(x)|^p dx \right)^{1/p} + \left(\int_{\Omega} |v(x)|^p dx \right)^{1/p} \tag{1.18}$$

1.3. Distribuciones

Para la elaboración de esta Sección, se ha seguido principalmente la referencia [1].

Recordemos que $\Omega \subset \mathbb{R}^N$ es un abierto no vacío.

Notación multiíndices de Schwart:

Si $u : \Omega \subset \mathbb{R}^N \longrightarrow \mathbb{R}$ es una función, utilizaremos la notación multiíndices de Schwart: $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_N) \in \mathbb{N}^N$ se llama multiíndice de dimensión N y $|\alpha| = \sum_{i=1}^N \alpha_i$ es la longitud de α . Si $\alpha, \beta \in \mathbb{N}^N$ son dos multiíndices, denotamos:

1. $\alpha! = \alpha_1! \alpha_2! \cdots \alpha_N!$
2. $\alpha \geq \beta$ si $\alpha_i \geq \beta_i$, $\forall i : 1 \leq i \leq N$.
3. $\alpha \geq \beta$, $\binom{\alpha}{\beta} = \frac{\alpha!}{\beta!(\alpha - \beta)!}$.
4. Si $x \in \mathbb{R}^N$ y $\alpha \in \mathbb{N}^N$, $x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_N^{\alpha_N}$.
5. Si $u : \Omega \subset \mathbb{R}^N \rightarrow \mathbb{R}$ es una función regular, denotamos

$$\partial^\alpha u(x) = \partial_1^{\alpha_1} \partial_2^{\alpha_2} \cdots \partial_N^{\alpha_N} u(x) = \frac{\partial^{|\alpha|} u}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \cdots \partial x_N^{\alpha_N}}(x); \quad \partial_i = \frac{\partial}{\partial x_i}$$

Sabemos

$$\begin{aligned} \mathcal{D}(\Omega) &\subset C_0^k(\Omega) \subset C^k(\Omega) \subset C^0(\Omega) \subset L_{Loc}^1(\Omega), \\ \mathcal{D}(\Omega) &\subset L^p(\Omega) \subset L_{Loc}^1(\Omega), \quad \forall p \in [1, \infty]. \end{aligned}$$

(Todos son espacios vectoriales, pero solo $L^p(\Omega)$ es un espacio de Banach).

Definición 1.32 Sean $\{v_n\}_{n \geq 1} \subset L_{Loc}^1(\Omega)$ y $v \in L_{Loc}^1(\Omega)$. Se dice que $v_n \rightarrow v$ en $L_{Loc}^1(\Omega)$ si $\int_K |v_n - v| dx \rightarrow 0$, $\forall K \subset \Omega$ compacto.

Definición 1.33 Sean $\{\varphi_n\}_{n \geq 1} \subset C^k(\Omega)$ y $\varphi \in C^k(\Omega)$ ($k \geq 0$). Se dice que $\varphi_n \rightarrow \varphi$ en $C^k(\Omega)$ si para todo $\alpha \in \mathbb{N}^N$ con $|\alpha| \leq k$,

$$\partial^\alpha \varphi_n \rightarrow \partial^\alpha \varphi \text{ uniformemente en } K, \quad \forall K \text{ compacto.}$$

Definición 1.34 Sean $\{\varphi_n\}_{n \geq 1} \subset C_0^k(\Omega)$ y $\varphi \in C_0^k(\Omega)$ ($k \geq 0$). Se dice que $\varphi_n \rightarrow \varphi$ en $C_0^k(\Omega)$ si existe un compacto $K \subset \Omega$ tal que

$$\text{sop } \varphi_n, \text{ sop } \varphi \subset K, \quad \forall n \geq 1$$

y para todo $\alpha \in \mathbb{N}^N$ con $|\alpha| \leq k$,

$$\partial^\alpha \varphi_n \rightarrow \partial^\alpha \varphi \text{ uniformemente en } K.$$

Definición 1.35 Sean $\{\varphi_n\}_{n \geq 1} \subset \mathcal{D}(\Omega)$ y $\varphi \in \mathcal{D}(\Omega)$. Se dice que $\varphi_n \rightarrow \varphi$ en $\mathcal{D}(\Omega)$ si existe un compacto $K \subset \Omega$ tal que

$$\text{sop } \varphi_n, \text{ sop } \varphi \subset K, \quad \forall n \geq 1$$

y para todo $\alpha \in \mathbb{N}^N$,

$$\partial^\alpha \varphi_n \rightarrow \partial^\alpha \varphi \text{ uniformemente en } K.$$

Observación 1.36

De las definiciones anteriores se tiene que si $\{\varphi_n\}_{n \geq 1} \subset \mathcal{D}(\Omega)$ y $\varphi \in \mathcal{D}(\Omega)$ son tales que $\varphi_n \rightarrow \varphi$ en $\mathcal{D}(\Omega)$, entonces $\varphi_n \rightarrow \varphi$ en $C_0^k(\Omega)$, $\forall k \geq 0$. Esta propiedad la escribiremos como

$$\mathcal{D}(\Omega) \hookrightarrow C_0^k(\Omega).$$

Diremos que $\mathcal{D}(\Omega)$ se inyecta en $C_0^k(\Omega)$ con inyección (secuencialmente) continua. Del mismo modo:

$$\begin{aligned} \mathcal{D}(\Omega) &\hookrightarrow C_0^k(\Omega) \hookrightarrow C^k(\Omega) \hookrightarrow C^0(\Omega) \hookrightarrow L_{Loc}^1(\Omega), \\ \mathcal{D}(\Omega) &\hookrightarrow L^p(\Omega) \hookrightarrow L_{Loc}^1(\Omega), \quad \forall p \in [1, \infty]. \end{aligned}$$

Definición 1.37 Una distribución sobre Ω es cualquier aplicación $T : \mathcal{D}(\Omega) \rightarrow \mathbb{R}$ lineal y secuencialmente continua, es decir, tal que si $\{\varphi_n\}_{n \geq 1} \subset \mathcal{D}(\Omega)$ y $\varphi \in \mathcal{D}(\Omega)$ satisfacen $\varphi_n \rightarrow \varphi$ en $\mathcal{D}(\Omega)$, entonces

$$T(\varphi_n) \rightarrow T(\varphi) \text{ en } \mathbb{R}.$$

Notación:

$$\mathcal{D}'(\Omega) = \{T : T \text{ es distribución sobre } \Omega\} \tag{1.19}$$

Si $T \in \mathcal{D}'(\Omega)$ y $\varphi \in \mathcal{D}(\Omega)$ escribiremos

$$\langle T, \varphi \rangle = T(\varphi) \in \mathbb{R}$$

Ejemplo 1.38 (La distribución de Dirac)

Fijemos $x \in \Omega$ con $\Omega \subset \mathbb{R}^N$ un abierto. Definimos δ_x mediante

$$\delta_x : \varphi \in \mathcal{D}(\Omega) \mapsto \langle \delta_x, \varphi \rangle = \varphi(x) \in \mathbb{R}.$$

Se tiene $\delta_x \in \mathcal{D}'(\Omega)$ pues es lineal y secuencialmente continua. Además δ_x no es la distribución nula, pues existe $\varphi \in \mathcal{D}(\Omega)$ tal que $\langle \delta_x, \varphi \rangle = \varphi(x) = 1$.

Definición 1.39 Dada una sucesión $\{T_n\}_{n \geq 1} \subset \mathcal{D}'(\Omega)$, diremos que T_n converge a un elemento $T \in \mathcal{D}'(\Omega)$ si para cada $\varphi \in \mathcal{D}(\Omega)$ se satisface $\langle T_n, \varphi \rangle \rightarrow \langle T, \varphi \rangle$ en \mathbb{R} , y escribiremos $T_n \rightarrow T$ en $\mathcal{D}'(\Omega)$.

Proposición 1.40 Dada $f \in L_{Loc}^1(\Omega)$, definimos

$$T_f : \varphi \in \mathcal{D}(\Omega) \mapsto \langle T_f, \varphi \rangle = \int_{\Omega} f\varphi \, dx \in \mathbb{R}.$$

Entonces, T_f está bien definida y $T_f \in \mathcal{D}'(\Omega)$. Además, el operador

$$\mathcal{T} : f \in L_{Loc}^1(\Omega) \mapsto \mathcal{T}(f) = T_f \in \mathcal{D}'(\Omega)$$

está bien definido, es lineal, inyectivo y secuencialmente continuo.

Observación 1.41

Mediante el operador T estamos identificando de forma biunívoca $L^1_{Loc}(\Omega)$ con un subconjunto de $\mathcal{D}'(\Omega)$. Como T es secuencialmente continuo, hay una inclusión de las topologías. Escribiremos

$$L^1_{Loc}(\Omega) \hookrightarrow \mathcal{D}'(\Omega)$$

Definición 1.42 Sean $T \in \mathcal{D}'(\Omega)$ y $\alpha \in \mathbb{N}^N$ un multiíndice. Se denomina derivada en el sentido de las distribuciones (o derivada distribucional) de orden α , $\partial^\alpha T$, como el elemento de $\mathcal{D}'(\Omega)$ dado por

$$\langle \partial^\alpha T, \varphi \rangle = (-1)^{|\alpha|} \langle T, \partial^\alpha \varphi \rangle, \quad \forall \varphi \in \mathcal{D}(\Omega). \tag{1.20}$$

$\partial^\alpha T$ está bien definida y $\partial^\alpha T \in \mathcal{D}'(\Omega)$ para todo $\alpha \in \mathbb{N}^N$.

Además, la aplicación $\partial^\alpha : T \in \mathcal{D}'(\Omega) \mapsto \partial^\alpha T \in \mathcal{D}'(\Omega)$ es lineal y secuencialmente continua.

El concepto de derivada distribucional es coherente con el concepto de derivada clásica:

Proposición 1.43 Sea $u \in C^0(\Omega)$ tal que $\exists \partial_i u(x), \forall x \in \Omega$, con $\partial_i u \in C^0(\Omega)$. Entonces $\partial_i(T_u) = T_{\partial_i u}$.

Veamos un ejemplo sobre la derivada distribucional.

Ejemplo 1.44

Calculamos la primera y segunda derivada en $\mathcal{D}'(\Omega)$ de

$$f(x) = |x| = \begin{cases} x & \text{si } x \geq 0 \\ -x & \text{si } x < 0 \end{cases}$$

En este caso tenemos $\Omega = \mathbb{R}$. Hay que comprobar que $f \in L^1_{Loc}(\Omega)$, es decir, que la integral sobre cada compacto es finita.

Sea $K \subset \mathbb{R}$ compacto. $\implies \max_K |x| = M \in (0, \infty)$. $\int_K f(x) dx \leq M|K| < \infty$

$f \in L^1_{Loc}(\Omega) \implies f \in \mathcal{D}'(\mathbb{R})$. Escriberemos $f (= T_f)$.

$$f' \in \mathcal{D}'(\mathbb{R}) \implies \langle f', \varphi \rangle = -\langle f, \varphi' \rangle = -\int_{\mathbb{R}} |x| \varphi'(x) dx \quad \forall \varphi \in \mathcal{D}(\mathbb{R}).$$

Sea $\varphi \in \mathcal{D}(\mathbb{R})$, entonces $\text{sop } \varphi \subset (-M, M)$.

$$\begin{aligned} \langle f', \varphi \rangle &= -\int_{-M}^M |x| \varphi'(x) dx = -\left(\int_{-M}^0 -x \varphi'(x) dx + \int_0^M x \varphi'(x) dx \right) \\ &= -\left(\int_{-M}^0 \varphi(x) dx - x\varphi(x)|_{-M}^0 - \int_0^M \varphi(x) dx + x\varphi(x)|_0^M \right) \\ &\stackrel{*}{=} -\left(\int_{-M}^0 \varphi(x) dx - \int_0^M \varphi(x) dx \right) = \int_{-M}^0 -\varphi(x) dx + \int_0^M \varphi(x) dx = \langle g, \varphi \rangle \\ &= \int_{\mathbb{R}} g(x) \varphi(x) dx \end{aligned}$$

$$\text{con } g(x) = \begin{cases} 1 & \text{si } x \geq 0 \\ -1 & \text{si } x < 0 \end{cases}$$

(*) : $\varphi(-M) = \varphi(M) = 0$ pues $\text{sop } \varphi \subset (-M, M)$.

$v = |x|$ es una función continua, no derivable (es derivable a trozos) y posee derivada distribucional.

Veamos ahora la segunda derivada.

$$\begin{aligned} \langle f'', \varphi \rangle &= -\langle f', \varphi' \rangle = \int_{-M}^0 \varphi'(x) dx - \int_0^M \varphi'(x) dx = \varphi(0) - \varphi(-M) - \varphi(M) + \varphi(0) \\ &= 2\varphi(0) = \langle 2\delta_0, \varphi \rangle. \end{aligned}$$

Por tanto, $f'' = 2\delta_0$.

1.4. Los espacios de Sobolev : $H^1(\Omega)$, $H_0^1(\Omega)$ y $H^{-1}(\Omega)$

El espacio $H^1(\Omega)$

Todos los resultados de este apartado se detallan en [7].

Definición 1.45 Sea $\Omega \subset \mathbb{R}^N$ un abierto no vacío. Definimos el espacio de Sobolev $H^1(\Omega)$ como:

$$H^1(\Omega) = \{u \in L^2(\Omega) : \frac{\partial u}{\partial x_i} = \partial_i u \in L^2(\Omega), \forall i : 1 \leq i \leq N\}. \quad (1.21)$$

Observación 1.46

Las derivadas consideradas son las derivadas distribucionales : $\frac{\partial u}{\partial x_i} \in L^2(\Omega)$ equivale a $\exists v_i \in L^2(\Omega)$ tal que

$$-\int_{\Omega} u(x) \frac{\partial \varphi}{\partial x_i}(x) dx = \int_{\Omega} v_i(x) \varphi(x) dx \quad \forall \varphi \in \mathcal{D}(\Omega). \quad (1.22)$$

Diremos en este caso que u es derivable débil (la derivada parcial de u en el sentido de las distribuciones es una función). Cada v_i se llama i -ésima derivada parcial débil de u y de ahora en adelante lo escribiremos como $\frac{\partial u}{\partial x_i}$.

Proposición 1.47 El espacio de Sobolev $H^1(\Omega)$ dotado del producto escalar

$$\langle u, v \rangle_{H^1(\Omega)} = \int_{\Omega} (u(x)v(x) + \nabla u(x) \cdot \nabla v(x)) dx \quad (1.23)$$

y la norma $\|u\|_{H^1(\Omega)} = \|u\|_{1,2;\Omega} = (\int_{\Omega} |u(x)|^2 + |\nabla u(x)|^2)^{1/2}$ es un espacio de Hilbert.

Además sabemos que en $H^1(\Omega)$ se cumple la fórmula de integración por partes.

Proposición 1.48 Sean $a, b \in \mathbb{R}, a < b$, entonces

$$\int_a^b u(x)v'(x) dx = u(b)v(b) - u(a)v(a) - \int_a^b u(x)'v(x) dx \quad \forall u, v \in H^1((a, b)) \quad (1.24)$$

En la práctica es importante saber si las funciones regulares son densas en $H^1(\Omega)$. Para ello vemos el siguiente resultado de densidad.

Teorema 1.49 Si Ω es un conjunto abierto acotado regular de clase C^1 , o bien $\Omega = \mathbb{R}_+^N$, o incluso si $\Omega = \mathbb{R}^N$, entonces $\mathcal{D}(\Omega)$ es denso en $H^1(\Omega)$.

El espacio $H_0^1(\Omega)$

Definamos ahora otro espacio de Sobolev que denotamos por $H_0^1(\Omega)$ y es un subespacio de $H^1(\Omega)$. Los siguientes resultados se detallan en [6].

Definición 1.50 $H_0^1(\Omega)$ es la adherencia de $\mathcal{D}(\Omega)$ en $H^1(\Omega)$, es decir, $H_0^1(\Omega) = \overline{\mathcal{D}(\Omega)}^{\|\cdot\|_{H^1(\Omega)}}$.

Proposición 1.51 El espacio de Sobolev $H_0^1(\Omega)$ dotado del producto escalar (1.23) (producto escalar de $H^1(\Omega)$) es un espacio de Hilbert.

Teorema 1.52 (Desigualdad de Poincaré) Sea $\Omega \subset \mathbb{R}^N$ abierto conexo acotado en alguna dirección. Entonces existe una constante $C = C(\Omega) > 0$ tal que

$$\int_{\Omega} |v(x)|^2 dx \leq C \int_{\Omega} |\nabla v(x)|^2 dx \quad \forall v \in H_0^1(\Omega) \quad (1.25)$$

Demostración: Lo hacemos utilizando un razonamiento de densidad. $\mathcal{D}(\Omega)$ es denso en $H_0^1(\Omega)$.

Consideramos Ω acotado en la dirección x_1 , $\Omega \subset (-M, M) \times \mathbb{R}^{N-1}$.

Sea $\varphi \in \mathcal{D}(\Omega) \implies \text{sop}(\varphi) \subset (-M, M) \times \mathbb{R}^{N-1} \implies \varphi(-M, x_2, \dots, x_N) = \varphi(M, x_2, \dots, x_N) = 0$.

$$x \in \Omega, \varphi(x) = \int_{-M}^{x_1} \frac{\partial}{\partial x_1} \varphi(y_1, x_2, \dots, x_N) dy_1$$

$$\begin{aligned} |\varphi(x)| &\leq \int_{-M}^{x_1} \left| \frac{\partial}{\partial x_1} \varphi(y_1, x_2, \dots, x_N) \right| dy_1 \leq \int_{-M}^M 1 \cdot \left| \frac{\partial}{\partial x_1} \varphi(y_1, x_2, \dots, x_N) \right| dy_1 \stackrel{\text{Hölder}}{\leq} \\ &\leq \left(\int_{-M}^M 1^2 dy \right)^{1/2} \cdot \left(\int_{-M}^M \left| \frac{\partial}{\partial x_1} \varphi(y_1, x_2, \dots, x_N) \right|^2 dy_1 \right)^{1/2} \end{aligned}$$

$$\implies |\varphi(x)|^2 \leq 2M \int_{-M}^M \left| \frac{\partial}{\partial x_1} \varphi(y_1, x_2, \dots, x_N) \right|^2 dy_1$$

$$\int_{\mathbb{R}^{N-1}} |\varphi(x)|^2 dx_2 \cdots dx_N \leq 2M \int_{-M}^M \int_{\mathbb{R}^{N-1}} \left| \frac{\partial \varphi}{\partial x_1}(x) \right|^2 dx$$

$$\int_{-M}^M \int_{\mathbb{R}^{N-1}} |\varphi(x)|^2 dx \leq (2M)^2 \int_{-M}^M \int_{\mathbb{R}^{N-1}} \left| \frac{\partial \varphi}{\partial x_1}(x) \right| dx \leq (2M)^2 \int_{-M}^M \int_{\mathbb{R}^{N-1}} |\nabla \varphi(x)|^2 dx \iff$$

$$\iff \int_{\Omega} |\varphi(x)|^2 dx \leq (2M)^2 \int_{\Omega} |\nabla \varphi(x)|^2 dx.$$

Sea $u \in H_0^1(\Omega) \implies \exists \{\varphi_n\}_{n \geq 1} \subset \mathcal{D}(\Omega)$ tal que $\varphi_n \rightarrow u$ en $H^1(\Omega)$, es decir, $\varphi_n \rightarrow u$ en $L^2(\Omega)$ y $\partial_i \varphi_n \rightarrow \partial_i u$ en $L^2(\Omega)$.

Por tanto tenemos $\int_{\Omega} |\varphi_n(x)|^2 dx \leq C(\Omega) \int_{\Omega} |\nabla \varphi_n(x)|^2 dx$.

Tomando límite en n :

$$\left. \begin{aligned} \int_{\Omega} |\varphi_n(x)|^2 dx &= \|\varphi_n\|_{2;\Omega}^2 \rightarrow \|u\|_{2;\Omega}^2 \\ \int_{\Omega} \left| \frac{\partial \varphi_n}{\partial x_i}(x) \right|^2 dx &= \|\partial_i \varphi_n\|_{2;\Omega}^2 \rightarrow \|\partial_i u\|_{2;\Omega}^2 \end{aligned} \right\} \implies \int_{\Omega} |u(x)|^2 dx \leq C(\Omega) \int_{\Omega} |\nabla u(x)|^2 dx$$

□

Nota 1.53 La demostración del Teorema 1.52 nos da un procedimiento constructivo, para calcular la constante $C = C(\Omega)$ en cada problema.

Observación 1.54

La *Desigualdad de Poincaré* no es cierta en el espacio $H^1(\Omega)$. Por ejemplo, tomamos Ω acotado y $v \equiv 1$. Observamos que $\nabla v \equiv 0$ y por tanto

$$\int_{\Omega} 1^2 dx = |\Omega| \not\leq C \cdot 0 = 0 \quad (|\Omega| \neq 0)$$

Gracias a la *Desigualdad de Poincaré* podemos definir una norma equivalente en $H_0^1(\Omega)$

Corolario 1.55 Sea $\Omega \subset \mathbb{R}^N$ abierto conexo acotado en alguna dirección. Entonces la seminorma

$$|v|_{H_0^1(\Omega)} = |v|_{1,2;\Omega} = \left(\int_{\Omega} |\nabla v(x)|^2 dx \right)^{1/2}$$

es una norma en $H_0^1(\Omega)$, que es equivalente a la inducida por $H^1(\Omega)$.

Por lo tanto concluimos que $(H_0^1(\Omega), |\cdot|_{1,2;\Omega})$ es un espacio de Hilbert.

El espacio $H^{-1}(\Omega)$

De forma genérica, sea H un espacio de Hilbert. Denotamos por H' a su dual, es decir, $H' = \mathcal{L}(H, \mathbb{R})$. Véase [7].

Definición 1.56 Denominaremos $H^{-1}(\Omega)$ al dual topológico de $H_0^1(\Omega)$, es decir, $H^{-1}(\Omega) = H'(H_0^1(\Omega)) = \mathcal{L}(H_0^1(\Omega), \mathbb{R})$.

Observación 1.57

1. $H^{-1}(\Omega)$ es el dual de un espacio de Hilbert. Por tanto, por el Teorema de Riesz, podríamos identificarlo (mediante el producto escalar de $H_0^1(\Omega)$) con $H_0^1(\Omega)$, aunque no lo haremos.

$H_0^1(\Omega)$ es un subespacio de $H^1(\Omega)$ y en particular un subespacio estricto de $L^2(\Omega)$.

Entonces, identificando $L^2(\Omega)$ con su dual, que es él mismo, tenemos :

$$H_0^1(\Omega) \subset L^2(\Omega) \equiv L^2(\Omega) \subset H^{-1}(\Omega)$$

2. Si $L \in H^{-1}(\Omega)$,

$$\|L\|_{H^{-1}(\Omega)} = \sup_{v \in H_0^1(\Omega), v \neq 0} \frac{\langle L, v \rangle}{\|v\|_{1,2;\Omega}}$$

Veamos ahora una caracterización para $H^{-1}(\Omega)$.

Proposición 1.58 El espacio $H^{-1}(\Omega)$ está caracterizado por

$$H^{-1}(\Omega) = \left\{ f = v_0 + \sum_{i=1}^N \frac{\partial v_i}{\partial x_i} \text{ con } v_0, v_1, \dots, v_N \in L^2(\Omega) \right\} \quad (1.26)$$

En otras palabras, toda forma lineal y continua $L : H_0^1(\Omega) \rightarrow \mathbb{R}, L \in H^{-1}(\Omega)$, se escribe $\forall \Phi \in H_0^1(\Omega)$

$$L(\Phi) = \int_{\Omega} \left(v_0 \Phi - \sum_{i=1}^N v_i \frac{\partial \Phi}{\partial x_i} \right) dx \text{ con } v_0, v_1, \dots, v_N \in L^2(\Omega)$$

1.5. Trazas y fórmulas de Green

Si $\Omega = \mathbb{R}_+^N = \{(x', x_N) \in \mathbb{R}^N : x_N > 0\} \equiv \mathbb{R}^{N-1} \times (0, +\infty)$ con $N \geq 2$, entonces $\partial \mathbb{R}_+^N = \mathbb{R}^{N-1} \times \{0\}$ y $L^2(\partial \Omega) = L^2(\mathbb{R}^{N-1})$.

En el caso en que Ω es un abierto acotado regular de clase C^1 , también es posible definir $L^2(\partial \Omega)$.

A la hora de definir los espacios $L^p(\Omega)$, en la Sección 1.2, vimos que no podemos hablar del valor de una función de $L^p(\Omega)$ en un punto de Ω . Por tanto tenemos que ver si ocurre lo mismo sobre $\partial \Omega$, ya que es un conjunto de medida nula. Para ellos utilizaremos el siguiente teorema.

Teorema 1.59 (De Trazas) Sea $\Omega \subseteq \mathbb{R}^N$ un abierto acotado regular de clase C^1 , o bien \mathbb{R}_+^N . Definimos el operador traza, γ_0

$$H^1(\Omega) \cap C(\bar{\Omega}) \rightarrow L^2(\partial \Omega) \cap C(\bar{\partial \Omega}) \quad (1.27)$$

$$v \mapsto v|_{\partial\Omega} = \gamma_0(v)$$

Este operador se puede extender por continuidad a $H^1(\Omega) \rightarrow L^2(\partial\Omega)$ y se denominará $\tilde{\gamma}_0$, que por abuso de notación lo llamaremos γ_0 . En particular $\|\gamma_0(v)\|_{L^2(\partial\Omega)} \leq C\|v\|_{H^1(\Omega)}$.

Teorema 1.60 (Fórmula de Green) Sea $\Omega \subset \mathbb{R}^N$ un abierto acotado regular de clase C^1 . Sean $u, v \in H^1(\Omega)$, entonces satisfacen:

$$\int_{\Omega} u(x) \frac{\partial v}{\partial x_i}(x) dx = - \int_{\Omega} v(x) \frac{\partial u}{\partial x_i}(x) dx + \int_{\partial\Omega} u(x)v(x)n_i(x) dS \quad (1.28)$$

donde $n = (n_i)_{1 \leq i \leq N}$, con n_i las componentes del vector normal exterior.

Notación: Por abuso del lenguaje, la mayoría de las veces, denotaremos $\gamma_0(v)$ por v .

Gracias al Teorema de Trazas 1.59 tenemos una caracterización para $H_0^1(\Omega)$.

Teorema 1.61 Sea $\Omega \subset \mathbb{R}^N$ un abierto acotado regular de clase C^1 . Entonces,

$$H_0^1(\Omega) = \ker(\gamma_0) = \{u \in H^1(\Omega) : \gamma_0(u) = 0\} \quad (1.29)$$

Proposición 1.62 Se tiene que $R(\gamma_0) = \text{Im}(\gamma_0)$ (rango o imagen de γ_0) es un subespacio vectorial propio de $L^2(\partial\Omega)$, que llamaremos espacio de trazas sobre la frontera.

$$R(\gamma_0) = H^{1/2}(\partial\Omega). \quad (1.30)$$

Podemos consultar los resultados de esta Sección en [1].

1.6. Resultado de compacidad

En un espacio de Hilbert de dimensión infinita, no es cierto que, de cada sucesión acotada se puede extraer una subsucesión convergente.

Definición 1.63 ■ Sea $T \in \mathcal{L}(X, Y)$. Se dice que T es compacto si la imagen de $T(B)$ de cualquier conjunto acotado $B \in X$ es relativamente compacta en Y .

- Si X es un espacio de Banach, $K \subset X$ compacto $\iff \forall \{x_n\}_{n \geq 1} \subset K$, existe una subsucesión $\{x_{n_k}\}_{k \geq 1}$ y $x \in K$ tal que $\{x_{n_k}\} \rightarrow x$ en X .
- Si X es un espacio de Banach, $K \subset X$ es relativamente compacto $\iff \forall \{x_n\}_{n \geq 1} \subset K$, existe una subsucesión $\{x_{n_k}\}_{k \geq 1}$ y $x \in X$ tal que $\{x_{n_k}\} \rightarrow x$ en X .

Denotamos por $\mathcal{K}(X, Y)$ el conjunto de operadores lineales, continuos y compactos de X en Y , es decir, $\mathcal{K}(X, Y) = \{T \in \mathcal{L}(X, Y) : T \text{ es compacto}\} \subset \mathcal{L}(X, Y)$. Es un subespacio vectorial de $\mathcal{L}(X, Y)$.

Si $Y = X$, entonces $\mathcal{K}(X) = \mathcal{K}(X, X)$.

Teorema 1.64 (Rellich-Kondrachov) *Si Ω es un abierto acotado regular de clase C^1 , entonces de toda sucesión acotada de $H^1(\Omega)$ se puede extraer una subsucesión convergente en $L^2(\Omega)$. Decimos que la aplicación inyección $i : v \in H^1(\Omega) \mapsto v \in L^2(\Omega)$ es compacta ($H^1(\Omega) \rightrightarrows L^2(\Omega)$).*

Lema 1.65 *La inyección $i_0 : H_0^1(\Omega) \rightarrow L^2(\Omega)$ es compacta.*

Para la elaboración de esta Sección, se ha seguido principalmente la referencia [1].

1.7. Otros espacios : $H^m(\Omega)$, $H(\text{div})$ y $W^{m,p}(\Omega)$

Los espacios $H^m(\Omega)$

Para este apartado, véase [7].

Definición 1.66 *De acuerdo con la notación multiíndice de Schwartz (Sección 1.3), sea $m \in \mathbb{Z}$, $m \geq 0$. Definimos el espacio de Sobolev*

$$H^m(\Omega) = \{v \in L^2(\Omega) : \forall \alpha, |\alpha| \leq m, \partial^\alpha v \in L^2(\Omega)\} \tag{1.31}$$

donde la derivada parcial $\partial^\alpha v$ es en el sentido débil.

Proposición 1.67 $H^m(\Omega)$ dotado del producto escalar

$$\langle u, v \rangle_{H^m(\Omega)} = \int_{\Omega} \sum_{|\alpha| \leq m} \partial^\alpha u(x) \partial^\alpha v(x) dx \tag{1.32}$$

y la norma $\|u\|_{H^m(\Omega)} = \sqrt{\langle u, u \rangle_{H^m(\Omega)}}$ es un espacio de Hilbert.

Teorema 1.68 *Sea Ω un abierto acotado regular de clase C^1 . Si $m > \frac{N}{2}$ entonces $H^m(\Omega)$ es un subespacio de $C(\overline{\Omega})$.*

Teorema 1.69 *Si Ω un abierto acotado regular de clase C^m , o bien $\Omega = \mathbb{R}_+^N$ entonces $\mathcal{D}(\overline{\Omega})$ es denso en $H^m(\Omega)$.*

Podemos generalizar el Teorema de Trazas 1.59 en el espacio $H^m(\Omega)$. Lo vemos para $m = 2$.

Teorema 1.70 *Sea Ω un abierto regular de clase C^1 . Definimos el operador de trazas γ_1*

$$H^2(\Omega) \cap C^1(\overline{\Omega}) \rightarrow L^2(\partial\Omega) \cap C(\overline{\partial\Omega}) \tag{1.33}$$

$$v \mapsto \gamma_1(v) = v \Big|_{\frac{\partial \Omega}{\partial \vec{n}}} = \nabla v \cdot \vec{n}$$

donde \vec{n} es el vector normal exterior a $\partial\Omega$. Esta aplicación se puede extender de manera continua a $H^2(\Omega) \rightarrow L^2(\partial\Omega)$ y se denominará $\tilde{\gamma}_1$, que por abuso de notación llamaremos γ_1 .

En particular, $\forall v \in H^2(\Omega)$ tenemos $\|\gamma_1(v)\| = \|\frac{\partial v}{\partial \vec{n}}\|_{H^2(\Omega)} \leq C\|v\|_{H^2(\Omega)}$.

Este teorema nos permite generalizarla *Fórmula de Green 1.60* en $H^2(\Omega)$.

Teorema 1.71 Sea Ω un abierto, regular de clase C^2 . Si $u \in H^2(\Omega)$ y $v \in H^1(\Omega)$, entonces

$$\int_{\Omega} \Delta u(x)v(x) dx = - \int_{\Omega} \nabla u(x) \cdot \nabla v(x) + \int_{\Omega} \frac{\partial u}{\partial n}(x)v(x) dS \quad (1.34)$$

El espacio $H(\text{div})$

Los siguientes conceptos los podemos encontrar en [1].

Definición 1.72 Definimos el espacio $H(\text{div})$ como

$$H(\text{div}) = \{w \in L^2(\Omega)^N : \nabla \cdot w \in L^2(\Omega)\}. \quad (1.35)$$

Proposición 1.73 $H(\text{div})$ dotado del producto escalar

$$\langle u, v \rangle_{H(\text{div})} = \int_{\Omega} u(x)v(x) dx + \int_{\Omega} (\nabla \cdot u(x))(\nabla \cdot v(x)) dx \quad (1.36)$$

y la norma $\|u\|_{H(\text{div})} = \sqrt{\langle u, u \rangle_{H(\text{div})}}$, es un espacio de Hilbert.

Teorema 1.74 Sea Ω un abierto acotado regular de clase C^1 , o bien $\Omega = \mathbb{R}_+^N$. Entonces se tiene:

1. $\mathcal{D}_{\overline{\Omega}}(\mathbb{R}^N)$ es denso en $H(\text{div})$.
2. $\exists \gamma_n \in \mathcal{L}(H(\text{div}), H^{-1/2}(\partial\Omega))$ tal que $\gamma_n(v) = v \cdot n|_{\partial\Omega}$ si $v \in H(\text{div}) \cap C^0(\overline{\Omega})^N$.
3. $\int_{\Omega} w(x)\nabla \cdot v(x) dx = - \int_{\Omega} \nabla w(x) \cdot v(x) dx + \langle \gamma_n(v), \gamma_0(w) \rangle_{H^{-1/2}, H^{1/2}} \quad \forall w \in H^1(\Omega), \forall v \in H(\text{div})$.

Los espacios $W^{m,p}(\Omega)$

Más generalmente, podemos definir los espacios $W^{m,p}(\Omega)$ para un entero $m \geq 0$ y un real $1 \leq p \leq +\infty$. Para más detalles, véase [6].

Definición 1.75 Para cada entero $m \geq 0$, definimos el espacio de Sobolev $W^{m,p}(\Omega)$ dado por

$$W^{m,p}(\Omega) = \{v \in L^p(\Omega) \text{ tal que } \forall \alpha \text{ con } |\alpha| \leq m, \partial^\alpha v \in L^p(\Omega)\}, \quad (1.37)$$

donde $\partial^\alpha v$ es la derivada parcial en el sentido débil.

Proposición 1.76 Considerando la norma

$$\|u\|_{W^{m,p}(\Omega)} = \left(\sum_{|\alpha| \leq m} \|\partial^\alpha u\|^p \right)^{1/p}$$

tenemos que $W^{m,p}(\Omega)$ son espacios de Banach.

Observación 1.77

Si consideramos $p = 2$,

$$W^{m,2}(\Omega) = H^m(\Omega).$$

Estos espacios son importantes para problemas no lineales y también para los lineales debido a las desigualdades de Sobolev.

Proposición 1.78 Sea Ω un abierto regular, o bien $\Omega = \mathbb{R}^N$, o bien $\Omega = \mathbb{R}_+^N$, entonces

$$\begin{cases} \text{Si } p < N & W^{1,p}(\Omega) \hookrightarrow L^q(\Omega) \quad \forall q \in [1, p^*] \text{ con } 1/p^* = 1/p - 1/N \\ \text{Si } p = N & W^{1,p}(\Omega) \hookrightarrow L^q(\Omega) \quad \forall q \in [1, +\infty) \\ \text{Si } p > N & W^{1,p}(\Omega) \hookrightarrow C(\bar{\Omega}) \end{cases} \quad (1.38)$$

1.8. Convergencia débil

A lo largo de esta sección X representa un espacio normado con norma que será denotada por $\|\cdot\|$. Recordemos que, dados una sucesión $\{x^n\}_{n \geq 1} \subset X$ y un elemento $x \in X$, se tiene que x^n converge hacia x en X (que denotaremos $x^n \rightarrow x$ en X) si la sucesión real $\|x^n - x\|$ converge hacia cero. En algunas ocasiones diremos que x^n converge en norma o fuertemente hacia x .

Veamos otro concepto de convergencia, [8].

Definición 1.79 Sean $\{x^n\}_{n \geq 1} \subset X$ una sucesión y $x \in X$ un elemento de X . Diremos que x^n converge débilmente hacia x si para cualquier $x' \in X'$ se tiene

$$\langle x', x^n \rangle_{X', X} \rightarrow \langle x', x \rangle_{X', X} \quad (1.39)$$

En esta definición $\langle \cdot, \cdot \rangle_{X', X}$ representa el producto de dualidad entre el espacio normado X y su dual X' .

Como su nombre indica, la convergencia débil es una idea “más débil” que la convergencia fuerte.

Proposición 1.80 Cualquier sucesión $\{x^n\}_{n \geq 1} \subset X$ fuertemente convergente es débilmente convergente.

Observación 1.81

Sin embargo, el recíproco es, en general, falso. En el caso de los espacios de Banach de dimensión finita, los dos conceptos coinciden.

Lema 1.82 *Sea $\{x^n\}_{n \geq 1}$ una sucesión débilmente convergente en un espacio normado X . Entonces:*

1. *El límite débil x de $\{x^n\}_{n \geq 1}$ es único.*
2. *Toda sucesión de $\{x^n\}_{n \geq 1}$ converge débilmente hacia x .*
3. *La sucesión $\{\|x^n\|\}_{n \geq 1}$ está acotada.*

Lema 1.83 *De toda sucesión $\{x^n\}_{n \geq 1}$ acotada en X podemos extraer una subsucesión que converge débilmente.*

Capítulo 2

Optimización

2.1. Introducción

La optimización puede considerarse como la búsqueda de la mejor solución (solución óptima) de un problema. El término mejor aquí, depende del contexto en el que se trabaje, podría significar solución que minimiza los costes, o maximiza los beneficios, o que hace que la distancia recorrida sea mínima, etc. Esta primera reflexión sobre lo que se entiende por optimización refleja claramente las importantísimas aplicaciones de esta área de las matemáticas a un amplio espectro de problemas.

El abordar un problema real de optimización supone básicamente dos etapas:

- Determinar el modelo matemático que rige el problema.
- Resolver dicho problema usando una serie de técnicas matemáticas.

En el siguiente esquema podemos observar las diferentes etapas o fases que hay que tener en cuenta a la hora de resolver dicho problema (ver figura 5.17).

En lo que sigue veremos problemas de optimización donde los modelos tienen naturaleza muy diferente. Se pueden clasificar según las características de las funciones que intervienen en ellos. Entre otros, el modelo puede ser una ecuación algebraica, puede depender de la solución de una ecuación diferencial ordinaria o bien de la solución de una ecuación en derivadas parciales. Para esta última será necesaria la teoría de control óptimo.

También podemos separar la optimización en dos grandes ramas, dependiendo si las variables son continuas o discretas.

Este capítulo considerará principalmente la cuestión de existencia y unicidad en optimización continua, ya sea en dimensión finita o infinita. En particular, veremos el crucial papel de la convexidad en la obtención de resultados de existencia en dimensión infinita.

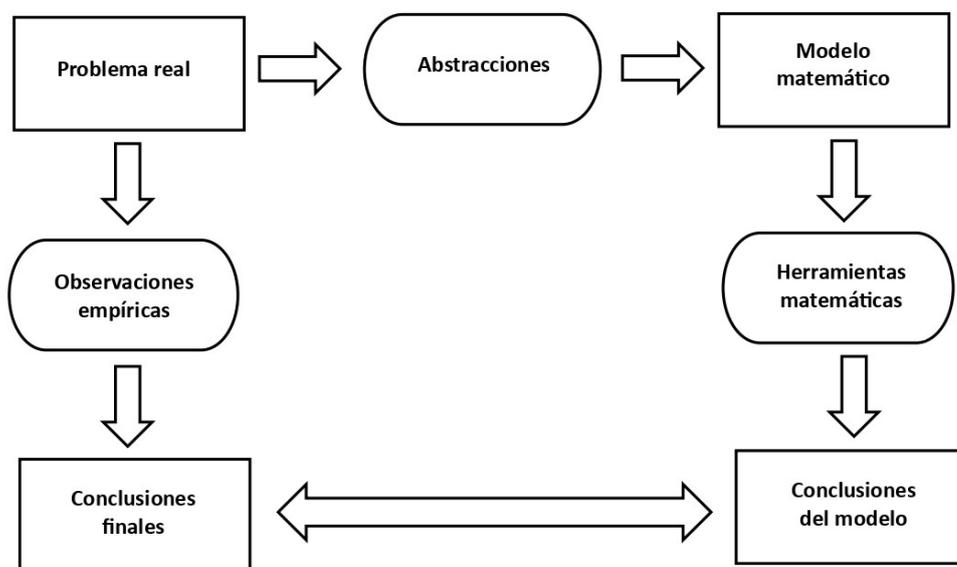


Figura 2.1: Fases del proceso de optimización

2.2. Ejemplos

En esta sección trataremos varios problemas típicos que nos permitirán explorar las diferentes “ramas” de la optimización. Para más ejemplos, véase [7].

2.2.1. En dimensión finita

Ejemplo 2.1 (Problema de transporte)

El problema de transporte es una de las primeras aplicaciones importantes de la programación lineal. Consiste en enviar unidades de un producto desde m orígenes, O_1, \dots, O_m , a n destinos, D_1, \dots, D_n , en las siguientes condiciones.

- Cada origen O_i , $i = 1, \dots, m$, dispone de una oferta a_i .
- Cada destino D_j , $j = 1, \dots, n$, realiza una demanda b_j .
- c_{ij} es el coste de enviar una unidad desde el origen O_i al destino D_j .

El problema es determinar el número de unidades x_{ij} que se deben enviar desde cada origen O_i hasta cada destino D_j para realizar el transporte a coste mínimo, teniendo en cuenta que hay que satisfacer las restricciones de oferta y demanda.

La formulación lineal de este problema es la siguiente:

$$\begin{aligned}
\min z &= \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij} \\
\text{sujeto a } &\sum_{i=1}^m x_{ij} = b_j \quad \forall j = 1, \dots, n \\
&\sum_{j=1}^n x_{ij} \leq a_j \quad \forall i = 1, \dots, m \\
&x_{ij} \geq 0 \quad \forall i = 1, \dots, m, \forall j = 1, \dots, n.
\end{aligned}$$

Las primeras n restricciones aseguran que se deben satisfacer las demandas de los destinos. Las m siguientes restricciones están asociadas a las ofertas de los orígenes, que no se deben sobrepasar. Las variables no pueden tomar valores negativos, ya que representan cantidades de producto que se transportan.

Ejemplo 2.2 (Problema de la mochila)

Consideremos que tenemos n objetos con pesos $p_1, \dots, p_n \in \mathbb{R}$ y utilidad respectiva $u_1, \dots, u_n \in \mathbb{R}$. Sea P el peso máximo que se puede llevar en la mochila.

Definimos las variables

$$x_i = \begin{cases} 1 & \text{si el objeto } i \text{ se mete en la mochila} \\ 0 & \text{en caso contrario} \end{cases}$$

Podemos formular el problema de la siguiente forma:

$$\begin{aligned}
\max z &= \sum_{i=1}^n u_i x_i \\
\text{sujeto a } &\sum_{i=1}^n p_i x_i \leq P \\
&x_i \in \{0, 1\} \quad \forall i = 1, \dots, n.
\end{aligned}$$

Ejemplo 2.3 (Problema del viajante de comercio)

Un ejemplo célebre de optimización combinatoria es el *Problema del viajante de comercio*, también conocido como *Traveling Salesman Problem (TSP)*. Consiste en lo siguiente:

“Un viajante quiere visitar n ciudades una y sólo una vez cada una, empezando por una cualquiera de ellas y regresando al mismo lugar del que partió. Supongamos que conoce la distancia entre cualquier par de ciudades. ¿De qué forma debe hacer el recorrido si pretende minimizar la distancia total?” ([9])

Podemos considerarlo como un grafo orientado de las n ciudades unidas por arcos ponderados por c_{ij} , siendo c_{ij} la distancia entre la ciudad i y la j . Debemos encontrar un ciclo en este grafo que pasa una sola vez por todas las ciudades.

2.2.2. En dimensión infinita

Para los siguientes ejemplos será necesario presentar una introducción a la teoría de control óptimo de sistemas gobernados por EDPs. Para más detalles véase [10].

Los dominios de aplicación de dicha teoría son múltiples: aeroespacial, automovilístico, robótico, aeronáutico, internet, comunicaciones en general, médico, químico, etc.

Para la formulación general de los problemas de control óptimo, se necesita una ecuación (o sistema) que debe ser complementada con condiciones de contorno y/o iniciales:

$$\begin{cases} E(x) = F(v) \\ CC/CI \end{cases} \quad (2.1)$$

En él, denotamos v un dato que debemos elegir: el *control*; y denotamos y la solución: el *estado*. También se necesita un conjunto U_{ad} de *controles admisibles* y una función J que a cada par (v, y) asigna un valor real. El objetivo consiste en minimizar J sujeto a que se tenga $v \in U_{ad}$ y se verifique (2.1). Se interpreta que queremos actuar sobre el sistema de manera que un coste se haga mínimo (o, equivalentemente, que un beneficio se haga máximo). Por motivos evidentes, los controles que minimizan J se suelen llamar *controles óptimos*.

Ejemplo 2.4 (comando óptimo)

Consideramos un sistema diferencial lineal con criterio cuadrático. El objetivo es guiar, por ejemplo, un robot, para que siga una trayectoria predefinida.

El estado del robot en el momento t está representado por una función $y(t)$ con valores en \mathbb{R}^N (típicamente, la posición y la velocidad). Actuamos sobre el robot mediante un comando $v(t)$ con valores en \mathbb{R}^M (la potencia del motor, la dirección de las ruedas, etc.). En la presencia de fuerzas $f(t) \in \mathbb{R}^N$ las leyes de la mecánica conducen a un sistema de ecuaciones diferenciales ordinarias (asumidas lineales por simplicidad).

$$\begin{cases} \frac{dy}{dt} = Ay + Bv + f, & 0 \leq t \leq T \\ y(0) = y_0 \end{cases} \quad (2.2)$$

donde $y_0 \in \mathbb{R}^N$ es el estado inicial del sistema, A y B son dos matrices constantes de dimensiones respectivas $N \times N$ y $N \times M$. Denotamos por $z(t)$ una trayectoria “objetivo” y z_T una posición “objetivo” final.

Para aproximar estos objetivos lo mejor posible y para minimizar el costo del control, introducimos tres matrices positivas simétricas R, Q, D donde solo R se supone además que es definida positiva. Entonces definimos un criterio cuadrático:

$$J(v) = \int_0^T Rv(t) \cdot v(t) dt + \int_0^T Q(y - z)(t) \cdot (y - z)(t) dt + D(y(T) - z_T) \cdot (y(T) - z_T).$$

Como los comandos admisibles son posiblemente limitados (la potencia de un motor a menudo está limitada . . .), introducimos un conjunto no vacío cerrado convexo K de \mathbb{R}^M que representa el conjunto de mandamientos admisibles. Por tanto, el problema es resolver

$$\inf_{v(t) \in K, 0 \leq t \leq T} J(v)$$

Será necesario, por supuesto, especificar en qué espacios de funciones minimizamos $J(v)$ y definimos la solución y de (2.2).

Ejemplo 2.5 (Control de una membrana)

Consideramos una membrana elástica, fija en su límite y deformada bajo la acción de una fuerza f . Este problema se modela por

$$\begin{cases} -\Delta u = f + v & \text{en } \Omega \\ u = 0 & \text{sobre } \partial\Omega \end{cases}$$

donde u es el desplazamiento vertical de la membrana y v es una fuerza de control disponible. Consideramos ω una parte del dominio Ω . Definimos el conjunto de controles admisibles

$$K = \{v(x) \text{ tal que } v_{\min} \leq v(x) \leq v_{\max} \text{ en } \Omega \text{ y } v = 0 \text{ en } \Omega \setminus \omega\}$$

donde v_{\min} y v_{\max} son dos funciones dadas. Buscamos el control que da el desplazamiento u lo más cerca posible de un desplazamiento deseado u_0 , y que tiene un costo moderado. Por lo tanto, definimos

$$J(v) = \frac{1}{2} \int_{\Omega} (|u - u_0|^2 + c|v|^2) dx$$

y escribiremos el problema de control (tendremos que especificar los espacios de funciones para los datos)

$$\inf_{v \in K} J(v)$$

2.3. Convexidad

Consideraremos principalmente algunos problemas de minimización, aunque basta cambiar el signo para obtener un problema de maximización. Más concretamente, para cierta función f y un conjunto X , se tiene

$$\sup_{x \in X} f(x) = \inf_{x \in X} -f(x)$$

El espacio en el que se encuentra nuestro problema será un espacio vectorial normado, V , dotado de la norma $\|\cdot\|$. $K \subset V$ será un subconjunto donde buscamos la solución, es el conjunto de elementos admisibles del problema (restricciones impuestas). Finalmente, la función objetivo a minimizar, la denotamos por J , está definida sobre K y toma valores en \mathbb{R} , es decir, $J: v \in K \rightarrow J(v) \in \mathbb{R}$.

Por tanto, el problema estudiado se denotará como

$$\inf_{v \in K \subset V} J(v) \tag{2.3}$$

Cuando usamos la notación ínf para un problema de optimización, esto indica que no se sabe a priori, si se alcanza el valor mínimo, es decir, si $\exists \bar{v} \in K$ tal que $J(\bar{v}) = \inf_{v \in K \subset V} J(v)$.

Si queremos indicar que se alcanza el valor mínimo, preferimos la notación

$$\min_{v \in K \subset V} J(v) \tag{2.4}$$

Lo mismo ocurre en los problemas de maximización reemplazando ínf y mín por sup y máx respectivamente.

Podemos consultar los resultados de esta Sección en [7] y [11].

Definición 2.6 Diremos que u es un mínimo local (o punto mínimo) de J sobre K si y solo si

$$u \in K \quad y \quad \exists \delta > 0 \text{ tal que } \forall v \in K \text{ con } \|v - u\| < \delta \text{ se tiene } J(u) \leq J(v). \tag{2.5}$$

Además, u será un mínimo global de J sobre K si y solo si

$$u \in K \quad y \quad J(u) \leq J(v) \quad \forall v \in K. \tag{2.6}$$

Definición 2.7 Diremos ínfimo de J sobre K (o más comúnmente, valor mínimo), que denotaremos por (2.3), para referirnos al límite superior en \mathbb{R} de las constantes que limitan por debajo a J en K . Si J no está acotado por debajo en K , entonces el ínfimo es $-\infty$. Si K es vacío, por convenio, el ínfimo es $+\infty$.

Llamaremos sucesión minimizante J sobre K a la sucesión $(u^n)_{n \geq 1}$ tal que

$$u^n \in K \quad \forall n \quad y \quad \lim_{n \rightarrow +\infty} J(u^n) = \inf_{v \in K} J(v) \tag{2.7}$$

Hay que diferenciar si nos encontramos en un espacio de dimensión finita o infinita, pues tienen distinto comportamiento. Resultados válidos para dimensión finita no son suficiente para dimensión infinita. Estudiamos a continuación ambos casos.

- Optimización en dimensión finita:

En este caso podemos suponer, sin pérdida de generalidad, que $V = \mathbb{R}^N$ con el producto escalar habitual $u \cdot v = \sum_{i=1}^N u_i v_i$ y la norma euclídea $\|u\| = \sqrt{u \cdot u}$.

El siguiente resultado nos garantiza la existencia de un mínimo.

Teorema 2.8 Sea K un conjunto cerrado no vacío de \mathbb{R}^N , y J una función continua sobre K que toma valores en \mathbb{R} y satisface la propiedad “infinito en infinito”, es decir,

$$\forall (u^n)_{n \geq 1} \text{ sucesión en } K \text{ tal que } \lim_{n \rightarrow \infty} \|u^n\| = +\infty \implies \lim_{n \rightarrow \infty} J(u^n) = +\infty \quad (2.8)$$

Entonces existe al menos un punto mínimo de J sobre K . Además en cada sucesión minimizante de J sobre K podemos extraer una subsucesión que converge a un punto mínimo sobre K .

Demostración: Sea $(u^n)_{n \geq 1}$ una sucesión minimizante de J sobre K . La condición (2.8) implica que $(u^n)_{n \geq 1}$ está acotada ya que $J(u^n)$ es una sucesión de reales acotados.

Por lo tanto, existe una subsucesión $(u^{n_k})_{k \geq 1}$ que converge a $u \in \mathbb{R}^N$. Como K es cerrado, $u \in K$, y por continuidad $J(u^{n_k})$ converge a $J(u)$. Finalmente, por la Definición 2.7 $J(u) = \inf_{v \in K} J(v)$.

□

▪ Optimización en dimensión infinita:

Como dijimos antes, el estudio en este caso es distinto. En dimensión infinita un conjunto acotado y cerrado no es compacto. Si seguimos trabajando en un espacio vectorial V con norma $\|\cdot\|$ (como en el caso finito) no podremos continuar al tener que usar la compacidad como en el Teorema 2.8. Podríamos seguir si se satisface la condición :

$$\begin{aligned} \forall (u^n)_{n \geq 1} \text{ en } K \text{ tal que } \sup_{n \in \mathbb{N}} \|u^n\| < \infty \text{ tenemos} \\ \lim_{n \rightarrow \infty} J(u^n) = l < \infty \implies \exists u \in K \text{ tal que } J(u) \leq l. \end{aligned} \quad (2.9)$$

No es verificable en general, pero si para los problemas de minimización convexa.

La convexidad desempeña un importante papel en la programación matemática, puesto que permite obtener resultados teóricos y aplicar métodos específicos.

Veremos el concepto de convexidad para conjuntos y para funciones, como también algunas caracterizaciones.

Dentro de la optimización matemática tienen una extremada importancia los problemas convexos, que son aquellos de la forma:

$$\min_{v \in K} J(v) \quad (2.10)$$

donde K es un conjunto convexo y J una función convexa.

En lo que sigue, supondremos que V es un espacio de Hilbert equipado con producto escalar denotado $\langle u, v \rangle$ y con norma asociada $\|v\|$.

Definición 2.9 K es un conjunto convexo si contiene todos los segmentos que unen dos de sus puntos, es decir, si

$$\forall x, y \in K \text{ y } \forall \lambda \in [0, 1] \text{ se verifica } \lambda x + (1 - \lambda)y \in K. \quad (2.11)$$

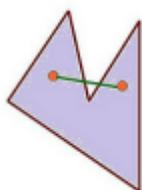


Figura 2.2: Conjunto no convexo

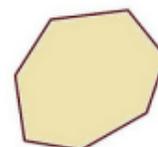


Figura 2.3: Conjunto convexo

Definición 2.10 Diremos que una función J definida sobre un conjunto convexo no vacío $K \subset V$ con valores en \mathbb{R} es convexa si y solo si

$$J(\theta u + (1 - \theta)v) \leq \theta J(u) + (1 - \theta)J(v) \quad \forall u, v \in K, \quad \forall \theta \in [0, 1]. \quad (2.12)$$

Además, J se dirá estrictamente convexa si la desigualdad (2.12) es estricta cuando $u \neq v$ y $\theta \in (0, 1)$.

Se dice que J es (estrictamente) cóncava si $-J$ es (estrictamente) convexa.

Definición 2.11 Si J es una aplicación definida sobre K con valores en \mathbb{R} , llamamos epígrafe de J al conjunto

$$Epi(J) = \{(\lambda, v) \in \mathbb{R} \times K, \lambda \geq J(v)\} \quad (2.13)$$

La demostración de la siguiente Proposición puede encontrarse en [11].

Proposición 2.12 J es convexo si y solo si $Epi(J)$ es un conjunto convexo de $\mathbb{R} \times V$.

En los casos que se plantean normalmente en la práctica, lo deseado siempre es encontrar soluciones globales, es decir, la mejor en términos absolutos. Pero, desafortunadamente, la mayor parte de las técnicas de optimización (principalmente las que están basadas en la derivabilidad) están encaminadas a la búsqueda de óptimos locales. Por esa razón, son necesarios resultados que permitan reconocer cuándo los óptimos locales son también globales.

El más importante de tales resultados es el teorema siguiente.

Teorema 2.13 Sea J es una función convexa en un conjunto convexo K . Entonces:

- i) Todo mínimo local de J sobre K es un mínimo global.
- ii) Además, si J es estrictamente convexa, entonces existe como máximo un punto mínimo.

Demostración:

- i) Sea u un mínimo local de J sobre K , de la Definición 2.6 tenemos que

$$\exists \delta > 0 \mid J(u) \leq J(w) \quad \forall w \in K \text{ tal que } \|w - u\| < \delta.$$

Veamos que $J(u) \leq J(v) \quad \forall v \in K$:

Tomamos $v \in K$. Para $\theta \in (0, 1)$ suficientemente pequeño $w_0 = \theta v + (1 - \theta)u$ satisface $\|w_0 - u\| < \delta$ y $w_0 \in K$, pues K es convexo. Por tanto, $J(u) \leq J(w_0)$.

La convexidad de J implica que $J(w_0) \leq \theta J(v) + (1 - \theta)J(u)$, por tanto $J(u) \leq \theta J(v) + (1 - \theta)J(u)$, que simplificando obtenemos $J(u) \leq J(v)$, es decir, u es un mínimo global sobre K .

ii) Por *i*) deducimos que J sólo tiene mínimos absolutos, pero por ser estrictamente convexa sólo tiene un mínimo absoluto. En efecto, supongamos que u_1, u_2 son dos mínimos absolutos, $u_1 \neq u_2$. Necesariamente se tendrá $J(u_1) = J(u_2)$.

Pero entonces para todo $\theta \in (0, 1)$ se tiene:

$$J(\theta u_2 + (1 - \theta)u_1) < \theta J(u_2) + (1 - \theta)J(u_1) = J(u_1) = J(u_2)$$

lo que contradice que u_1 y u_2 son mínimos absolutos.

□

A partir de ahora, utilizaremos la idea de “convexidad fuerte” en vez de la convexidad estricta.

Definición 2.14 Sea V un espacio de Hilbert y $K \subset V$ un subconjunto convexo no vacío. Diremos que una función $J : K \subset V \rightarrow \mathbb{R}$ es fuertemente convexa en K si J es continua en K y $\exists \alpha > 0$ tal que, para cualesquiera $u, v \in K$ y $\theta \in [0, 1]$, se tiene:

$$J((1 - \theta)u + \theta v) \leq (1 - \theta)J(u) + \theta J(v) - \alpha \frac{\theta(1 - \theta)}{2} \|u - v\|^2 \quad (2.14)$$

También se dice que J es α -convexa o que es un funcional elíptico.

Para la Proposición 2.16 necesitaremos previamente el siguiente resultado.

Teorema 2.15 (Separación de un punto y un conjunto convexo) Sean K un conjunto cerrado no vacío convexo de un espacio de Hilbert V y $x_0 \notin K$. Entonces existe un hiperplano cerrado de V que separa estrictamente x_0 y K , es decir, existe una forma lineal $L \in V'$ y $\alpha \in \mathbb{R}$ tales que

$$L(x_0) < \alpha < L(x) \quad \forall x \in K. \quad (2.15)$$

Demostración: Denotemos por x_K la proyección de x_0 sobre K . Como $x_0 \notin K$, tenemos $x_K - x_0 \neq 0$. Sea L la forma lineal definida para todo $y \in V$ por $L(y) = \langle x_K - x_0, y \rangle$ y sea $\alpha = \frac{L(x_K) + L(x_0)}{2}$. De (1.8) tenemos $L(x) \geq L(x_K) > \alpha > L(x_0) \quad \forall x \in K$, lo que finaliza la prueba.

□

Proposición 2.16 Si J es convexa y continua sobre un conjunto no vacío cerrado y convexo K , entonces existe una forma lineal $L \in V'$ y una constante $\delta \in \mathbb{R}$ tal que

$$J(v) \geq L(v) + \delta \quad \forall v \in K. \quad (2.16)$$

Si además J es fuertemente convexa sobre K , entonces existen dos constantes $\gamma > 0$ y $\delta \in \mathbb{R}$ tal que

$$J(v) \geq \gamma \|v\|^2 - \delta \quad \forall v \in K. \quad (2.17)$$

Demostración: Como J es continua y convexa sobre K , se tiene que $Epi(J)$ es un subconjunto cerrado (pues J es continuo sobre K), convexo (pues J es convexo en U) y no vacío. Tomemos $v_0 \in K$ y $\lambda_0 \in \mathbb{R}$ tales que $\lambda_0 \leq J(v_0)$. Como $(\lambda_0, v_0) \notin Epi(J)$, podemos aplicar el Teorema de separación de un punto y un convexo, deduciendo la existencia de $\alpha, \beta \in \mathbb{R}$ y una forma lineal $L \in V'$ tales que

$$\beta \lambda + L(v) > \alpha > \beta \lambda_0 + L(v_0) \quad \forall (\lambda, v) \in Epi(J). \quad (2.18)$$

Como, para v fija, podemos tomar λ arbitrariamente grande en el lado izquierdo de (2.18), es evidente que $\beta \geq 0$. Además, como podemos tomar $v = v_0$ en el lado izquierdo de (2.18), β no puede ser cero.

Por tanto, tenemos $\beta > 0$ y deducimos que $J(v) + \frac{L(v)}{\beta} > \frac{\alpha}{\beta} \quad \forall v \in K$, lo que prueba (2.16).

Veamos ahora (2.17).

Consideramos $v_0 \in K$ fijo. Para $v \in K$ (2.14) y (2.16) implican

$$\frac{J(v)}{2} + \frac{J(v_0)}{2} \geq J\left(\frac{v+v_0}{2}\right) + \frac{\alpha}{8} \|v - v_0\|^2 \geq \frac{L(v) + L(v_0)}{2} + \frac{\alpha}{8} \|v - v_0\|^2 + \delta.$$

De aquí deducimos

$$J(v) \geq \frac{\alpha}{4} \|v\|^2 - \frac{\alpha}{2} \langle v, v_0 \rangle + L(v) + C_1,$$

con $C_1 = \frac{\alpha}{4} \|v_0\|^2 + L(v_0) - J(v_0) + 2\delta$. Aplicando la desigualdad de Cauchy-Schwarz, deducimos

$$J(v) \geq \frac{\alpha}{4} \|v\|^2 - \left(\|L\|_{V'} + \frac{\alpha \|v_0\|}{2} \right) \|v\| + C_1 \geq \frac{\alpha}{8} \|v\|^2 - C,$$

para cierta constante C .

□

Para los resultados de existencia y unicidad V será un espacio de Hilbert, K un conjunto cerrado convexo de V y J una función continua convexa sobre K .

Veamos un primer resultado en el caso que J es fuertemente convexa.

Teorema 2.17 *Sea K un conjunto no vacío cerrado convexo de un espacio de Hilbert V y J una función continua α -convexa sobre K . Entonces, existe un único mínimo u de J sobre K . Además, se cumple*

$$\|v - u\|^2 \leq \frac{4}{\alpha} (J(v) - J(u)) \quad \forall v \in K. \quad (2.19)$$

Demostración: Sea $(u^n)_{n \geq 1}$ una sucesión minimizante de J sobre K . De (2.17) deducimos que J está acotada inferiormente sobre K , y para $n, m \in \mathbb{N}$, (2.14), tomando $\theta = \frac{1}{2}$, implica que

$$\frac{\alpha}{8} \|u^n - u^m\|^2 + J\left(\frac{u^n + u^m}{2}\right) - \inf_{v \in K} J(v) \leq \frac{1}{2} \left(J(u^n) - \inf_{v \in K} J(v) \right) + \frac{1}{2} \left(J(u^m) - \inf_{v \in K} J(v) \right).$$

Esto demuestra que la sucesión $(u^n)_{n \geq 1}$ es de Cauchy y por tanto converge a u que es un punto mínimo de J sobre K , ya que J es continua y K cerrado.

La unicidad la hemos probado en el Teorema 2.13.

Finalmente, si consideramos $v \in K$, como K es convexo $\frac{u+v}{2} \in K$, y aplicando (2.14) con $\theta = \frac{1}{2}$ obtenemos

$$\frac{\alpha}{8} \|u - v\|^2 \leq \frac{J(u)}{2} + \frac{J(v)}{2} - J\left(\frac{u+v}{2}\right) \leq \frac{J(u) - J(v)}{2},$$

ya que $J(u) \leq J\left(\frac{u+v}{2}\right)$.

□

Es posible generalizar el Teorema 2.17 al caso de funciones J que son sólo convexas (y no fuertemente convexas).

Teorema 2.18 *Sea K un conjunto convexo cerrado no vacío de un espacio de Banach reflexivo V y J una función continua convexa sobre K y que satisface la propiedad “infinito en infinito” (2.8). Entonces existe un mínimo de J sobre K .*

Demostración: Sea $(u^n)_{n \geq 1}$ una sucesión minimizante de J sobre K , de (2.8) deducimos que $(u^n)_{n \geq 1}$ está acotada. Gracias al Lema 1.82, podemos considerar una subsucesión $(u^{n_k})_{k \geq 1}$ que converge débilmente al límite $u \in V$. Por ser J una función continua y convexa se satisface

$$J(u) \leq \lim_k \inf J(u^{n_k}) = \inf_{v \in K} J(v),$$

de donde se concluye que u es un mínimo de J sobre K .

□

Observación 2.19

El Teorema 2.18 nos garantiza la existencia de un mínimo, pero no dice nada acerca de la unicidad o la estimación del error.

Capítulo 3

Optimización diferenciable

La Teoría de la diferenciación en espacios de dimensión infinita tiene sus comienzos en 1887 con V. Volterra, al considerar éstas las derivadas variacionales de funciones de $\mathcal{C}([a, b])$ en \mathbb{R} . Una de las más importantes motivaciones para el desarrollo del Cálculo Diferencial en espacios de dimensión infinita proviene de la teoría clásica del Cálculo de Variaciones.

A partir de ahora, supondremos que V es un espacio real de Hilbert y que J es una función continua con valores en \mathbb{R} . El producto escalar en V se denota $\langle u, v \rangle$ y la norma asociada $\|u\|$.

3.1. La diferencial de Fréchet y de Gâteaux

Para la elaboración de esta Sección, se han seguido principalmente [7] y [13].

Definición 3.1 Diremos que la función J , definida en un entorno de $u \in V$ con valores en \mathbb{R} , es diferenciable en el sentido de Fréchet en u si existe una forma lineal continua en V , $L \in V'$, tal que

$$J(u + w) = J(u) + L(w) + o(w), \quad \text{con } \lim_{w \rightarrow 0} \frac{|o(w)|}{\|w\|} = 0 \quad (3.1)$$

Llamamos a L la derivada de J en u y denotamos $L = J'(u)$.

Observación 3.2

La Definición 3.1 es de hecho válida si V es sólo un espacio de Banach (no usamos el producto escalar en (3.1)).

Podemos generalizar la definición anterior en el caso de que el espacio de llegada sea otro espacio de Hilbert W .

Definición 3.3 La función f , definida en un entorno de $u \in V$ con valores en W , es diferenciable en el sentido de Fréchet en u si existe una aplicación lineal y continua, $L : V \rightarrow W$, tal que

$$f(u + w) = f(u) + L(w) + o(w), \quad \text{con } \lim_{w \rightarrow 0} \frac{\|o(w)\|}{\|w\|} = 0 \quad (3.2)$$

Ejemplo 3.4

Sea $(X, \|\cdot\|)$ un espacio normado, para la función real $J : X \rightarrow \mathbb{R}$ definida por $J(x) = \|x\|^2$, se tiene que J es diferenciable en el origen y su diferencial es la aplicación nula en X .

En efecto,

$$\lim_{w \rightarrow 0} \frac{J(0+w) - J(0) - L(w)}{\|w\|} = \lim_{w \rightarrow 0} \frac{\|w\|^2}{\|w\|} = 0$$

considerando L como la aplicación nula en X .

Existen otras ideas de diferenciabilidad, más débiles que en el *sentido de Fréchet*.

Definición 3.5 Diremos que la función J , definida en un entorno de $u \in V$ con valores en \mathbb{R} , es diferenciable en el sentido de Gâteaux en u y en la dirección $w \in V$ si existe el siguiente límite:

$$\lim_{\delta \rightarrow 0} \frac{J(u + \delta w) - J(u)}{\delta} \quad (3.3)$$

Diremos que es diferenciable en el sentido de Gâteaux en $u \in V$ si el límite (3.3) existe $\forall w \in V$.

Vemos a continuación dos ejemplos en los cuales en uno se cumple que la función es diferenciable en el *sentido de Gâteaux* y en el otro no.

Ejemplo 3.6

Consideremos el espacio de Banach $X := \mathcal{C}([a, b])$ con su norma usual y consideremos el funcional $J : X \rightarrow \mathbb{R}$ definido por

$$J(u) = \int_a^b (\operatorname{sen}^3(x) + u^2(x)) dx$$

el cual está definido en todo X . Veamos ahora la variación de Gâteaux en una dirección w y en un punto u . Tenemos que evaluar:

$$\begin{aligned} \frac{J(u + \delta w) - J(u)}{\delta} &= \frac{1}{\delta} \int_a^b ((u + \delta w)^2(x) - u^2(x)) dx \\ &= \frac{1}{\delta} \int_a^b (u^2(x) + 2\delta w(x)u(x) + \delta^2 w^2(x) - u^2(x)) dx \\ &= 2 \int_a^b u(x)w(x) dx + \delta \int_a^b w^2(x) dx \end{aligned}$$

Con lo cual se tiene que

$$\lim_{\delta \rightarrow 0} \frac{J(u + \delta w) - J(u)}{\delta} = 2 \int_a^b u(x)w(x) dx$$

Esto significa que existe la derivada de Gâteaux en toda dirección y además es una aplicación lineal y continua por lo tanto es Gâteaux diferenciable.

Ejemplo 3.7

Sea $(X, \|\cdot\|)$ un espacio normado y $J : X \rightarrow \mathbb{R}$ definida por $J(x) = \|x\|$. Si $w \in X$ es no nulo, y tomamos $u = 0$ obtenemos el siguiente límite

$$\lim_{\delta \rightarrow 0} \frac{|\delta| \|w\|}{\delta}$$

Es evidente que no existe, por lo tanto el funcional F no tiene derivada de Gâteaux en ninguna dirección en el punto $u = 0$.

Proposición 3.8 Sea $F : V \rightarrow \mathbb{R}$ una función diferenciable en el sentido de Fréchet en u_0 entonces, F es diferenciable en el sentido de Gâteaux en ese punto. ([12])

Demostración: Sea $w \in V$ un vector no nulo. En efecto, como F es diferenciable Fréchet sabemos que $L \in V'$ con lo cual fijado w se tendrá que

$$\lim_{\delta \rightarrow 0} \frac{F(u_0 + \delta w) - F(u_0) - L(u_0)(\delta w)}{\|\delta w\|} = 0$$

lo que implica que

$$\lim_{\delta \rightarrow 0} \frac{|\delta|}{\delta} \frac{F(u_0 + \delta w) - F(u_0) - \delta L(u_0)(w)}{|\delta| \|w\|} = 0$$

y simplificando

$$\lim_{\delta \rightarrow 0} \left(\frac{F(u_0 + \delta w) - F(u_0)}{\delta} - L(u_0)(w) \right) = 0$$

Por tanto hemos probado que si F es diferenciable en el sentido de Fréchet entonces también lo es en el sentido de Gâteaux.

□

Veamos un ejemplo que muestra que la implicación contraria es falsa.

Ejemplo 3.9

$$F(x, y) = \begin{cases} \frac{xy^3}{x^2 + y^6} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{si } (x, y) = (0, 0) \end{cases}$$

Esta función no es continua en el punto $(0, 0)$ y entonces no puede ser diferenciable Fréchet en dicho punto. Sin embargo, las derivadas direccionales son todas nulas, y por lo tanto la función es diferenciable Gâteaux en el origen.

Proposición 3.10 Sea J una aplicación diferenciable Gâteaux de V en \mathbb{R} y $\alpha > 0$. Las siguientes afirmaciones son equivalentes

$$J \text{ es } \alpha\text{-convexa sobre } V \tag{3.4}$$

$$J(v) \geq J(u) + \langle J'(u), v - u \rangle + \frac{\alpha}{2} \|v - u\|^2, \quad \forall u, v \in V \tag{3.5}$$

$$\langle J'(u) - J'(v), u - v \rangle \geq \alpha \|u - v\|^2, \quad \forall u, v \in V \tag{3.6}$$

Demostración: Veamos en primero lugar (3.4) \implies (3.5).

Suponemos que J es α -convexa sobre V , entonces aplicando (2.14) con $\theta \in (0, 1)$, obtenemos

$$\frac{J(u + \theta(v - u)) - J(u)}{\theta} \leq J(v) - J(u) - \frac{\alpha}{2}(1 - \theta)\|v - u\|^2.$$

Tomando límite cuando $\theta \rightarrow 0^+$, deducimos (3.5).

Veamos ahora (3.5) \implies (3.6).

Consideramos (3.5), $u, v \in V$ y lo aplicamos a (u, v) , (v, u) y sumamos:

$$J(v) \geq J(u) + \langle J'(u), v - u \rangle + \frac{\alpha}{2}\|v - u\|^2, \quad J(u) \geq J(v) + \langle J'(v), u - v \rangle + \frac{\alpha}{2}\|u - v\|^2,$$

entonces $-\langle J'(u), v - u \rangle - \langle J'(v), u - v \rangle \geq \alpha\|u - v\|^2$, de donde deducimos

$$\langle J'(u) - J'(v), u - v \rangle \geq \alpha\|u - v\|^2.$$

Por último, (3.6) \implies (3.4).

Para $u, v \in V$, $t \in \mathbb{R}$, consideramos $\varphi(t) = J(u + t(v - u))$. Entonces φ es derivable y por lo tanto continua en \mathbb{R} , $\varphi'(t) = \langle J'(u + t(v - u)), v - u \rangle$.

De (3.6), $\varphi'(t) - \varphi'(s) = \alpha(t - s)\|v - u\|^2$, si $t \geq s$. Tomando $\theta \in (0, 1)$ e integrando respecto de t en el intervalo $[\theta, 1]$ y respecto de s en el intervalo $[0, \theta]$, obtenemos

$$\theta\varphi(1) + (1 - \theta)\varphi(0) - \varphi(\theta) \geq \frac{\alpha\theta(1 - \theta)}{2}\|v - u\|^2,$$

es decir, J es α -convexa sobre V .

□

Observación 3.11

Un caso particular de la proposición anterior es el caso $\alpha = 0$, en el que J es convexa sobre V .

Proposición 3.12 Sea J una función diferenciable Gâteaux sobre V . Será estrictamente convexa si y solo si

$$J(w) > J(v) + \langle J'(v), w - v \rangle, \quad \forall v, w \in V, v \neq w. \quad (3.7)$$

Demostración: Sean $v, w \in V$, dos puntos distintos de V y $\mu \in (0, 1)$. Sea $\lambda \in (0, 1)$ un número fijo. Como

$$v + \mu(w - v) = \left(1 - \frac{\mu}{\lambda}\right)v + \frac{\mu}{\lambda}(v + \lambda(w - v)),$$

se deduce:

$$J(v + \mu(w - v)) \leq \left(1 - \frac{\mu}{\lambda}\right)J(v) + \frac{\mu}{\lambda}J(v + \lambda(w - v)), \quad 0 \leq \mu \leq \lambda.$$

Por tanto, si la función J es estrictamente convexa, puesto que $\lambda < 1$ se tiene:

$$\frac{J(v + \mu(w - v)) - J(v)}{\mu} \leq \frac{J(v + \mu(w - v)) - J(v)}{\lambda} < J(w) - J(v), \quad 0 \leq \mu \leq \lambda$$

Ahora,

$$J'(v)(w - v) = \lim_{\mu \rightarrow 0} \frac{J(v + \mu(w - v)) - J(v)}{\mu} < J(w) - J(v).$$

Recíprocamente, supongamos que

$$J(w) > J(v) + \langle J'(v), w - v \rangle, \quad \forall v, w \in V.$$

Sean $v, w \in V$ dos puntos distintos de V y $\mu \in (0, 1)$. Se tiene, entonces, en particular:

$$\begin{aligned} J(w) &> J(w + \mu(v - w)) - \mu J'(w + \mu(v - w))(v - w) \\ J(v) &> J(w + \mu(v - w)) + (1 - \mu) J'(w + \mu(v - w))(v - w) \end{aligned}$$

Basta sumar las dos desigualdades multiplicadas por $(1 - \mu)$ y μ respectivamente para obtener:

$$J(\mu v + (1 - \mu)w) < \mu J(v) + (1 - \mu)J(w),$$

que establece la estricta convexidad de J .

□

Definición 3.13 Sea J una función de V en \mathbb{R} . Decimos que J es dos veces diferenciable Gâteaux en $u \in V$ si J es diferenciable Gâteaux en un entorno de u y si su derivada $J'(u)$ es diferenciable Gâteaux en u . Denotamos por $J''(u)$ la segunda derivada de J en u que satisface

$$J'(u + w) = J'(u) + J''(u)w + o(w) \text{ con } \lim_{w \rightarrow 0} \frac{\|o(w)\|}{\|w\|} = 0 \quad (3.8)$$

Lema 3.14 Si J es una función dos veces diferenciable Gâteaux de V en \mathbb{R} , entonces satisface

$$J(u + w) = J(u) + J'(u)w + \frac{1}{2}J''(u)(w, w) + o(\|w\|^2) \text{ con } \lim_{w \rightarrow 0} \frac{o(\|w\|^2)}{\|w\|^2} = 0. \quad (3.9)$$

Proposición 3.15 Si J es una función dos veces diferenciable Gâteaux sobre V , entonces las condiciones de la Proposición 3.10 son equivalentes a

$$J''(u)(w, w) \geq \alpha \|w\|^2 \quad \forall u, w \in V. \quad (3.10)$$

Ejercicio 3.16 Sea $a(\cdot, \cdot)$ una forma bilineal simétrica y continua sobre $V \times V$. Sea L una forma lineal continua sobre V . Tomamos $J(u) = \frac{1}{2}a(u, u) - L(u)$. Entonces, J es diferenciable sobre V y la derivada en el sentido de Fréchet es $a(u, w) - L(w)$, $\forall u, w \in V$. Además J es dos veces diferenciable sobre V y $J''(u)(v, w) = a(v, w)$.

Demostración: Aplicamos la Definición 3.1. Tenemos:

$$J(u + w) = \frac{1}{2}a(u + w, u + w) - L(u + w) = \frac{1}{2}(a(u, u) + a(w, w) + 2a(u, w)) - L(u) - L(w)$$

Entonces,

$$J(u + w) - J(u) = \frac{1}{2}a(w, w) + a(u, w) - L(w).$$

Considerando $o(w) = \frac{1}{2}a(w, w)$ y aplicando la continuidad de la forma bilineal:

$$0 \leq \lim_{w \rightarrow 0} \frac{|\frac{1}{2}a(w, w)|}{\|w\|} \leq \lim_{w \rightarrow 0} \frac{\frac{1}{2}M\|w\|^2}{\|w\|} = 0 \implies \lim_{w \rightarrow 0} \frac{|o(w)|}{\|w\|} = 0.$$

Por tanto,

$$J'(u)(w) = a(u, w) - L(w).$$

Calculamos ahora $J'(u+v)(w) = a(u+v, w) - L(w) = a(u, w) + a(v, w) - L(w)$. Entonces,

$$J'(u+v)(w) - J'(u)(w) = a(u, w) + a(v, w) - L(w) - a(u, w) + L(w) = a(v, w)$$

Consideramos $o(w) = 0$, $\lim_{w \rightarrow 0} \frac{\|o(w)\|}{\|w\|} = 0 \implies J''(u)(v, w) = a(v, w)$.

□

3.2. Condiciones de optimalidad (caso convexo)

Para la elaboración de esta Sección, se han seguido principalmente [7] y [13].

El siguiente resultado da una condición necesaria de mínimo local de un funcional.

Teorema 3.17 *Sea V un espacio normado, $\Omega \subset V$ un subconjunto y $J : \Omega \rightarrow \mathbb{R}$ un funcional. Supongamos que J alcanza un mínimo local en $u \in \Omega$ y que J es diferenciable Gâteaux en u . Entonces, se verifica la ecuación de Euler*

$$J'(u) = 0. \tag{3.11}$$

Demostración: Sea $v \in V$. Para t suficientemente pequeño se tendrá, por la condición de mínimo local

$$J(u + tv) \geq J(u).$$

Por otra parte, como J es diferenciable Gâteaux en u , por definición se tiene

$$J'(u)(v) = \langle J'(u), v \rangle = \lim_{t \rightarrow 0} \frac{J(u + tv) - J(u)}{t} = \begin{cases} \lim_{t \rightarrow 0^+} \frac{J(u + tv) - J(u)}{t} \geq 0 \\ \lim_{t \rightarrow 0^-} \frac{J(u + tv) - J(u)}{t} \leq 0 \end{cases}$$

Esto prueba que $\langle J'(u), v \rangle \geq 0 \quad \forall v \in V$, es decir, $J'(u) = 0$.

□

Asumimos que el conjunto de restricciones K es cerrado no vacío y convexo, y que J es un funcional continuo sobre un conjunto abierto contenido en K .

Teorema 3.18 (Desigualdad de Euler) *Sea $u \in K$ con K convexo. Suponemos que J es diferenciable Gâteaux en u . Si u es un punto mínimo local de J sobre K entonces*

$$\langle J'(u), v - u \rangle \geq 0, \quad \forall v \in K. \tag{3.12}$$

Si $u \in K$ satisface (3.12) y si J es convexo, entonces u es un mínimo global de J sobre K .

Demostración: Como u es un mínimo local de J sobre K sabemos que

$$J(u) \leq J(v) \text{ para } u \in K \text{ tal que } \exists \delta > 0 : \|v - u\| < \delta \quad \forall v \in K.$$

Sea $v \in K$. Por ser K convexo se tiene

$$\forall \lambda \in [0, 1], \quad u + \lambda(v - u) = \lambda v + (1 - \lambda)u \in K,$$

y además,

$$\langle J'(u), v - u \rangle = \lim_{\lambda \rightarrow 0} \frac{J(u + \lambda(v - u)) - J(u)}{\lambda} \geq 0$$

pues $J(u + \lambda(v - u)) \geq J(u)$ para λ suficientemente pequeño.

Si $u \in K$ satisface (3.12) y si J es convexo, entonces aplicando (3.5) con $\alpha = 0$ tenemos:

$$J(v) \geq J(u) + \langle J'(u), v - u \rangle \geq J(u),$$

de donde, por la Definición 2.6, concluimos que u es un mínimo global de J sobre K . □

Esta es una condición necesaria para la optimalidad que se vuelve necesaria y suficiente si J es convexo.

Observación 3.19

1. Un caso particular es considerar que K es un subespacio afín cerrado de V , es decir, $K = u_0 + \mathcal{P}$, con $u_0 \in V$ y \mathcal{P} un subespacio vectorial cerrado de V . Entonces, (3.12) es equivalente a

$$\langle J'(u), w \rangle = 0 \quad \forall w \in \mathcal{P}.$$

2. Supongamos ahora que K es un cono convexo cerrado, lo que significa que K es un conjunto cerrado convexo tal que $\lambda v \in K$ para todo $v \in K$ y todo $\lambda \geq 0$. Tomando $v = 0$ entonces $v = 2u$ en (3.12), obtenemos

$$\langle J'(u), u \rangle = 0$$

y

$$\langle J'(u), w \rangle \geq 0, \quad \forall w \in K.$$

Veamos ahora una condición de optimalidad de segundo orden. Para ello nos hará falta el siguiente teorema.

Proposición 3.20 Suponemos que $K = V$ y que J es dos veces diferenciable en u . Si u es un punto mínimo local de J , entonces

$$J'(u) = 0 \quad \text{y} \quad J''(u)(w, w) \geq 0, \quad \forall w \in V. \tag{3.13}$$

Por el contrario, si, para todo v en un entorno de u ,

$$J'(u) = 0 \quad \text{y} \quad J''(v)(w, w) \geq 0, \quad \forall w \in V. \tag{3.14}$$

entonces u es un mínimo local de J .

Demostración: Si u es un punto mínimo local, ya sabemos que se cumple $J'(u) = 0$ y por tanto de (3.9) deducimos (3.13).

Por el contrario, si u satisface (3.14), escribimos la Expresión de Taylor de segundo orden (en un entorno de cero) con resto exacto para la función $\Phi(t) = J(u + tw)$ con $t \in \mathbb{R}$ y deducimos que u es un mínimo local de J por la Definición 2.6.

□

Capítulo 4

Métodos numéricos en optimización sin restricciones

Sea J una función que es α -convexa y diferenciable definida sobre el espacio real de Hilbert V . Consideramos el problema sin restricción

$$\inf_{v \in V} J(v) \quad (4.1)$$

Trabajaremos con los siguientes metodos que nos permitirán aproximar la solución de (4.1), todos son iterativos.

4.1. Algoritmo de gradiente con paso óptimo

Los *métodos de descenso con paso óptimo* son métodos de descenso en los que el parámetro ρ^n es elegido para conseguir el máximo descenso en la dirección elegida, es decir:

$$\text{Dado } u^0 \in V, \quad u^{n+1} = u^n + \rho^n d^n, \quad n \geq 0$$

donde ρ^n es tal que

$$J(u^n + \rho^n d^n) = \inf_{\rho \in \mathbb{R}} J(u^n + \rho d^n).$$

El método de descenso con paso óptimo más conocido es el *método de gradiente con paso óptimo*, que corresponde a tomar $d^n = -J'(u^n)$.

Este algoritmo se puede encontrar en [13].

Algoritmo:

1. Elegir $u^0 \in V$.
2. Dados $n \geq 0$ y $u^n \in V$: calcular $J'(u^n) \in V$.

3. Calcular $\rho^n \in \mathbb{R}$ tal que $J(u^n - \rho^n J'(u^n)) = \inf_{\rho \in \mathbb{R}} J(u^n - \rho J'(u^n))$.
4. Tomar $u^{n+1} = u^n - \rho^n J'(u^n)$, hacer $n = n + 1$ y volver al paso 2.

Estudieemos la convergencia de dicho algoritmo.

Teorema 4.1 *Supongamos que J es α -convexa y diferenciable y que J' es Lipschitziana sobre todos los conjuntos acotados de V , es decir que*

$$\forall M > 0, \exists C_M > 0, \quad \|v\|, \|w\| \leq M \implies \|J'(v) - J'(w)\| \leq C_M \|v - w\|. \quad (4.2)$$

Entonces el algoritmo de gradiente con paso óptimo converge: para cualquier u^0 , la sucesión (u^n) dada por dicho algoritmo converge a la solución u de (4.1).

Demostración: Sin pérdida de generalidad supondremos que $J'(u^n) \neq 0 \quad \forall n \geq 0$. Pues si $J'(u^n) = 0$ para algún n , entonces $u^n = u$ y el método converge.

Haremos la demostración en varias etapas:

1. La aplicación $\phi^n : \rho \in \mathbb{R} \longrightarrow \phi^n(\rho) = J(u^n - \rho J'(u^n))$ es continua, estrictamente convexa y coercitiva.

En efecto, utilizando el hecho de que J es estrictamente convexa, $\forall \lambda \in (0, 1)$ con $\rho^1 \neq \rho^2$, se tiene:

$$\begin{aligned} \phi^n(\lambda \rho^1 + (1 - \lambda) \rho^2) &= J(u^n - (\lambda \rho^1 + (1 - \lambda) \rho^2) J'(u^n)) \\ &= J(\lambda(u^n - \rho^1 J'(u^n)) + (1 - \lambda)(u^n - \rho^2 J'(u^n))) \\ &< \lambda J(u^n - \rho^1 J'(u^n)) + (1 - \lambda) J(u^n - \rho^2 J'(u^n)) \\ &= \lambda \phi^n(\rho^1) + (1 - \lambda) \phi^n(\rho^2) \end{aligned}$$

Como J es α -convexa por la Proposición 3.10 se tiene

$$J(v) \geq J(w) + \langle J'(w), v - w \rangle + \frac{\alpha}{2} \|v - w\|^2 \quad \forall v, w \in V,$$

es decir:

$$\phi^n(\rho) \geq J(u^n) - \rho \|J'(u^n)\|^2 + \frac{\alpha}{2} \rho^2 \|J'(u^n)\|^2 = J(u^n) + \left(\frac{\alpha}{2} \rho^2 - \rho \right) \|J'(u^n)\|^2,$$

de donde

$$\lim_{|\rho| \rightarrow \infty} \phi^n(\rho) = +\infty,$$

lo que implica que ϕ^n es coercitiva.

2. Dos direcciones de descenso consecutivas son ortogonales:

$$\langle J'(u^{n+1}), J'(u^n) \rangle = 0$$

En efecto, puesto que $\phi^n : \mathbb{R} \rightarrow \mathbb{R}$ es continua, estrictamente convexa y coercitiva se deduce que tiene un único mínimo ρ^n en \mathbb{R} ,

$$\phi^n(\rho^n) = \inf_{\rho \in \mathbb{R}} \phi^n(\rho)$$

es decir:

$$J(u^n - \rho^n J'(u^n)) = \inf_{\rho \in \mathbb{R}} J(u^n - \rho J'(u^n)).$$

Por tanto ρ^n es el paso óptimo buscado y

$$u^{n+1} = u^n - \rho^n J'(u^n).$$

Se tendrá que $(\phi^n)'(\rho^n) = 0$ y entonces

$$0 = (\phi^n)'(\rho^n) = -J'(u^n)' J'(u^n - \rho^n J'(u^n)) = -\langle J'(u^{n+1}), J'(u^n) \rangle$$

3. La sucesión $\{J(u^n)\}_{n \geq 0}$ es monótona estrictamente decreciente.

En efecto, dado $u^{n+1} = u^n - \rho^n J'(u^n)$ tenemos $u^{n+1} - u^n = -\rho^n J'(u^n)$ y por tanto $\langle J'(u^{n+1}), u^{n+1} - u^n \rangle = 0$. Utilizando de nuevo la Proposición 3.10 :

$$\begin{aligned} J(u^n) &\geq J(u^{n+1}) + \langle J'(u^{n+1}), u^n - u^{n+1} \rangle + \frac{\alpha}{2} \|u^n - u^{n+1}\|^2 \\ &= J(u^{n+1}) + \frac{\alpha}{2} \|u^n - u^{n+1}\|^2 \end{aligned} \quad (4.3)$$

4.

$$\lim_{n \rightarrow \infty} \|u^n - u^{n+1}\| = 0 \text{ y } \lim_{n \rightarrow \infty} \|J'(u^n) - J'(u^{n+1})\| = 0$$

La sucesión $\{J(u^n)\}_{n \geq 0}$ es monótona decreciente y acotada inferiormente por $J(u)$. Por tanto, es convergente y el primer límite se concluye de la desigualdad (4.3). Por ser $\{J(u^n)\}_{n \geq 0}$ monótona decreciente y J coercitiva se deduce que $\{u^n\}_{n \geq 0}$ es acotada. Además, $J'(\cdot)$ es continua (J es α -convexa) y por tanto uniformemente continua sobre compactos. De ello y de $\lim_{n \rightarrow \infty} \|u^n - u^{n+1}\| = 0$ se deduce que $\lim_{n \rightarrow \infty} \|J'(u^n) - J'(u^{n+1})\| = 0$.

5.

$$\lim_{n \rightarrow \infty} J'(u^n) = 0.$$

En primer lugar debido a la ortogonalidad de dos direcciones consecutivas se tiene:

$$\|J'(u^n)\|^2 = \langle J'(u^n), J'(u^n) - J'(u^{n+1}) \rangle \leq \|J'(u^n)\| \|J'(u^n) - J'(u^{n+1})\|$$

y por tanto,

$$\|J'(u^n)\| \leq \|J'(u^n) - J'(u^{n+1})\|.$$

El resultado se concluye de 4.

6.

$$\lim_{n \rightarrow \infty} u^n = u \text{ y además } \|u^n - u\| \leq \frac{1}{\alpha} \|J'(u^n)\| \quad (4.4)$$

$$\alpha \|u^n - u\|^2 \leq \langle J'(u^n) - J'(u), u^n - u \rangle = \langle J'(u^n), u^n - u \rangle \leq \|J'(u^n)\| \|u^n - u\|,$$

de donde se concluyen ambos resultados al aplicar 5.

□

4.2. Algoritmo de gradiente con paso fijo

Se trata de un método más simple que el de *gradiente con paso óptimo* ya que en cada etapa n en lugar de calcular ρ^n resolviendo un problema de mínimos unidimensional, consideramos ρ un parámetro positivo fijo. Por tanto, la sucesión (u^n) viene dada por

$$u^{n+1} = u^n - \rho J'(u^n) \quad (4.5)$$

Este algoritmo se puede encontrar en [7].

Teorema 4.2 *Suponemos que J es α -convexa y diferenciable y que J es Lipschitziana sobre V , es decir que existe una constante $C > 0$ tal que*

$$\|J'(v) - J'(w)\| \leq C \|v - w\|, \quad \forall v, w \in V. \quad (4.6)$$

Entonces, si $0 < \rho < \frac{2\alpha}{C^2}$, el algoritmo de gradiente con paso fijo converge: para cualquier u^0 , la sucesión (u^n) definida por $u^{n+1} = u^n - \rho J'(u^n)$ converge a la solución u de (4.1).

Demostación : Tomamos $v^n = u^n - u$. Bajo las condiciones del enunciado se tiene que el funcional J admite un único mínimo $u \in V$ que satisface la ecuación de Euler $J'(u) = 0$ en V . Por tanto, $v^{n+1} = v^n - \rho(J'(u^n) - J'(u))$, de donde deducimos lo siguiente:

$$\begin{aligned} \|v^{n+1}\|^2 &= \|v^n\|^2 - 2\rho \langle J'(u^n) - J'(u), v^n \rangle + \rho^2 \|J'(u^n) - J'(u)\|^2 \\ &\leq \|v^n\|^2 - 2\rho \langle J'(u^n) - J'(u), v^n \rangle + \rho^2 C^2 \|v^n\|^2 \\ &\leq \|v^n\|^2 - 2\rho \alpha \|v^n\|^2 + \rho^2 C^2 \|v^n\|^2 \\ &= (1 - 2\rho \alpha + \rho^2 C^2) \|v^n\|^2 \end{aligned}$$

Consideramos ahora el polinomio $P(\rho) = 1 - 2\rho \alpha + \rho^2 C^2$.

Calculamos $P'(\rho) = -2\alpha + 2\rho C^2$ y $P(0) = 1$.

$$\min_{\rho \geq 0} P(\rho) = P\left(\frac{\alpha}{C^2}\right) = 1 - \frac{\alpha^2}{C^2}, \text{ veamos cuando se cumple } 0 < 1 - 2\rho \alpha + \rho^2 C^2 < 1$$

Si $\rho(C^2\rho - 2\alpha) < 0 \implies \rho < 0$ o $C^2\rho - 2\alpha < 0$. No es posible $\rho < 0$ pues por estar en el *Algoritmo de gradiente con paso fijo* sabemos que $\rho \leq 0$, por tanto tenemos $C^2\rho - 2\alpha < 0$. De ahí deducimos $\rho < \frac{2\alpha}{C^2}$. Concluimos que para $0 < \rho < \frac{2\alpha}{C^2}$ tenemos $1 - 2\rho\alpha + \rho^2 C^2 \in (0, 1)$ y por tanto deducimos la convergencia. □

En la práctica para elegir el algoritmo con el que trabajaremos utilizaremos la *tasa de convergencia* del algoritmo, que fija el número de iteraciones necesarias que hay que hacer para que el error $u^n - u$ sea menor que una tolerancia fijada a priori.

4.3. Método del gradiente conjugado para funcionales cuadráticos

Los métodos anteriores utilizaban en cada iteración una dirección de descenso localmente óptima ligada al funcional J considerado (la de $J'(u^n)$). Sin embargo, es posible encontrar direcciones de descenso mejoradas con respecto a la del gradiente.

Se trata de métodos iterativos que aplicados al caso de un funcional cuadrático $J : \mathbb{R}^N \rightarrow \mathbb{R}$ obtienen el óptimo en un número de pasos menor o igual que N .

Este método se puede encontrar en [13].

Veamos primero una breve introducción sobre los *métodos de direcciones conjugadas*.

Sea $J : \mathbb{R}^N \rightarrow \mathbb{R}$ un funcional cuadrático cualquiera,

$$J(v) = \frac{1}{2} \langle Av, v \rangle - \langle b, v \rangle = \frac{1}{2} v' Av - v' b,$$

donde, $A \in \mathcal{M}_{N \times N}$ es una matriz simétrica, definida positiva y $b \in \mathbb{R}^N$.

El principio de los métodos de direcciones conjugadas consiste en a partir de un punto $u^0 \in \mathbb{R}^N$, minimizar $J(v)$ siguiendo sucesivamente N direcciones linealmente independientes d^0, d^1, \dots, d^{N-1} que poseen la propiedad de ser *mutuamente conjugadas con respecto a la forma cuadrática $J(v)$* , es decir,

$$(d^i)' A d^j = 0, \quad i \neq j, \quad 0 \leq i, j \leq N - 1.$$

Supongamos entonces que para $n = 0, \dots, N - 2$, u^{n+1} se calcula a partir de u^n de la forma:

$$u^{n+1} = u^n + \rho^n d^n, \quad J(u^n + \rho^n d^n) = \min_{\rho \in \mathbb{R}} J(u^n + \rho d^n).$$

Entonces el punto u^N obtenido en la etapa N , es decir:

$$u^N = u^{N-1} + \rho^{N-1} d^{N-1} = \dots = u^0 + \sum_{j=0}^{N-1} \rho^j d^j,$$

es necesariamente el óptimo del problema, es decir, verifica:

$$J'(u^N) = Au^N - b = 0$$

Como ρ^k minimiza la función $\phi^n(\rho) = J(u^n + \rho d^n)$, se tendrá, $(\phi^n)'(\rho^n) = 0$, es decir:

$$(d^n)'J'(u^{n+1}) = (d^n)'J'(u^n + \rho^n d^n) = (d^n)'(A(u^n + \rho^n d^n) - b)$$

Por tanto, $(d^n)'(Au^n - b) + \rho^n(d^n)'Ad^n = 0$, de donde deducimos

$$\rho^n = -\frac{(d^n)'(Au^n - b)}{(d^n)'Ad^n}$$

Obsérvese que el denominador en la fórmula anterior no se anula por ser la matriz A definida positiva.

Teniendo en cuenta que

$$u^n = u^0 + \sum_{j=0}^{n-1} \rho^j d^j,$$

se puede escribir

$$(d^n)'Au^n = (d^n)'Au^0 + \sum_{j=0}^{n-1} \rho^j (d^n)'Ad^j = (d^n)'Au^0$$

y por tanto tenemos equivalentemente

$$\rho^n = -\frac{(d^n)'(Au^0 - b)}{(d^n)'Ad^n}$$

Ahora podemos probar la siguiente propiedad, característica de los *métodos de direcciones conjugadas*.

Teorema 4.3 *Para todo n ($1 \leq n \leq N$) el punto:*

$$u^n = u^0 + \sum_{j=0}^{n-1} \rho^j d^j,$$

es el óptimo de J en la variedad afín $V_n = u^0 + \langle d^0, d^1, \dots, d^{n-1} \rangle$. En particular,

$$u^N = u^0 + \sum_{j=0}^{N-1} \rho^j d^j,$$

es el óptimo de J en \mathbb{R}^N .

Demostración: Observamos que para todo i ($0 \leq i \leq n-1$)

$$(d^i)'Au^n = (d^i)'Au^0 + \sum_{j=0}^{n-1} \rho^j (d^i)'Ad^j = (d^i)'Au^0 + \rho^i (d^i)'Ad^i,$$

de donde se deduce, utilizando la expresión encontrada anteriormente para ρ^i :

$(d^i)'Au^n = (d^i)'Au^0 - (d^i)'(Au^0 - b)$ y finalmente $(d^i)'(Au^n - b) = 0$, $i = 0, 1, \dots, n-1$.

Esto demuestra que $J'(u^n) = Au^n - b$ es ortogonal a la variedad lineal V_n , es decir, $J'(u^n)(v - w) = 0$, $\forall v, w \in V_n$, y por tanto u^n es el óptimo de J en V_n .

En particular, para $n = N$, $V_N = \mathbb{R}^N$ (pues d^0, d^1, \dots, d^{N-1} son N vectores linealmente independientes) y u^N es el mínimo de J en \mathbb{R}^N .

□

Así pues, un *método de direcciones conjugadas* converge en N etapas, a lo sumo, cuando se trata de minimizar un funcional cuadrático en \mathbb{R}^N .

Como existen diversas posibilidades de elección de direcciones conjugadas d^0, d^1, \dots, d^{N-1} , se pueden construir numerosos algoritmos basados en el principio general de direcciones conjugadas.

Veamos el *método del gradiente conjugado*.

Consideramos $J : \mathbb{R}^N \rightarrow \mathbb{R}$ un funcional cuadrático cualquiera,

$$J(v) = \frac{1}{2}\langle Av, v \rangle - \langle b, v \rangle = \frac{1}{2}v'Av - v'b,$$

donde, $A \in \mathcal{M}_{N \times N}$ es una matriz simétrica, definida positiva y $b \in \mathbb{R}^N$.

La idea del *método del gradiente conjugado*, introducido por Hestenes y Stiefel en 1952, es construir progresivamente las direcciones d^0, d^1, \dots, d^{N-1} , mutuamente conjugadas con respecto a la matriz A . En cada etapa n , la dirección d^n es obtenida por combinación lineal de $-J'(u^n)$ y de las direcciones precedentes d^0, d^1, \dots, d^{n-1} ,

$$d^n = -J'(u^n) + \sum_{j=0}^{n-1} \beta^j d^j = -(Au^n - b) + \sum_{j=0}^{n-1} \beta^j d^j,$$

eligiendo los coeficientes β^j de la combinación lineal para que d^n sea conjugada con las direcciones anteriores. Notamos por $g_n := J'(u^n) = Au^n - b$. Entonces el método toma la forma siguiente

$$u^0 \in \mathbb{R}^N \text{ dado. Sean } g_0 = Au^0 - b, \quad d^0 = -g_0,$$

$$\begin{cases} u^{n+1} = u^n + \rho^n d^n, \text{ con } \rho^n = -\frac{\langle d^n, g_n \rangle}{\langle Ad^n, d^n \rangle} & n \geq 0 \\ d^{n+1} = -g_{n+1} + \beta^n d^n, \text{ con } \beta^n = \frac{\langle Ad^n, g_{n+1} \rangle}{\langle Ad^n, d^n \rangle} & n \geq 0 \end{cases}$$

Para demostrar que el algoritmo converge en N etapas es suficiente verificar que las direcciones d^0, d^1, \dots, d^{N-1} , son mutuamente conjugadas con respecto a la matriz A .

Teorema 4.4 *En una iteración n cualquiera del algoritmo en la que el óptimo aun no se haya alcanzado, es decir, $J'(u^i) = Au^i - b \neq 0$, $i = 0, 1, \dots, n$, se tiene*

$$\begin{aligned}\rho^n &= \frac{\langle g_n, g_n \rangle}{\langle Ad^n, d^n \rangle} \neq 0, \\ \beta^n &= \frac{\langle g_{n+1} - g_n, g_{n+1} \rangle}{\langle g_n, g_n \rangle} = \frac{\langle g_{n+1}, g_{n+1} \rangle}{\langle g_n, g_n \rangle}.\end{aligned}$$

Además, las direcciones d^0, d^1, \dots, d^{n+1} generadas por el algoritmo son mutuamente conjugadas.

Demostración: Se razona por inducción en n suponiendo que d^0, d^1, \dots, d^n son conjugadas.

De la definición del algoritmo se tiene

$$\begin{aligned}d^n &= -g_n + \beta^{n-1}d^{n-1}, \\ \rho^n &= -\frac{\langle d^n, g_n \rangle}{\langle Ad^n, d^n \rangle} = \frac{\langle g_n, g_n \rangle}{\langle Ad^n, d^n \rangle} - \beta^n \frac{\langle g_n, d_{n-1} \rangle}{\langle Ad^n, d^n \rangle}\end{aligned}$$

Como d^0, d^1, \dots, d^{n-1} son conjugadas, u^n es el mínimo de J sobre la variedad $V_n = u^0 + \langle d^0, d^1, \dots, d^{n-1} \rangle$. Por tanto, $\langle g_n, d^{n-1} \rangle = 0$ y se obtiene la expresión deseada para ρ^n .

Notemos ahora que

$$g_{n+1} - g_n = A(u^{n+1} - u^n) = \rho^n Ad^n.$$

Se tiene entonces que

$$\langle Ad^n, g_{n+1} \rangle = \frac{1}{\rho^n} \langle g_{n+1} - g_n, g_{n+1} \rangle.$$

Utilizando la expresión que acabamos de obtener para ρ^n , podemos poner:

$$\begin{aligned}\beta^n &= \frac{\langle Ad^n, g_{n+1} \rangle}{\langle Ad^n, d^n \rangle} = \frac{\langle Ad^n, d^n \rangle}{\langle g_n, g_n \rangle} \cdot \frac{\langle g_{n+1} - g_n, g_{n+1} \rangle}{\langle Ad^n, d^n \rangle} \\ &= \frac{\langle g_{n+1} - g_n, g_{n+1} \rangle}{\langle g_n, g_n \rangle} = \frac{\langle g_{n+1}, g_{n+1} \rangle}{\langle g_n, g_n \rangle}\end{aligned}$$

donde la última igualdad es consecuencia de que

$$g_n = d^n - \beta^{n-1}d^{n-1} \in \langle d^0, d^1, \dots, d^n \rangle$$

y g_{n+1} es ortogonal a este subespacio.

Probamos ahora que d^{n+1} es una dirección conjugada con d^0, d^1, \dots, d^n : $\langle Ad^i, d^{n+1} \rangle = 0$, $i = 0, 1, \dots, n$. En primer lugar se tiene por la definición de d^{n+1} y β^n :

$$\langle Ad^n, d^{n+1} \rangle = \langle Ad^n, -g_{n+1} + \beta^n d^n \rangle = -\langle Ad^n, g_{n+1} \rangle + \beta^n \langle Ad^n, d^n \rangle = 0.$$

Falta verificar que $\langle Ad^i, d^{n+1} \rangle = 0$, $i = 0, 1, \dots, n-1$. Se tiene por la hipótesis de inducción:

$$\langle Ad^i, d^{n+1} \rangle = -\langle Ad^i, g_{n+1} \rangle + \beta^n \langle Ad^i, d^n \rangle = -\langle Ad^i, g_{n+1} \rangle.$$

Puesto que $u^{i+1} = u^i + \rho^i d^i$ y $\rho^i \neq 0$ tenemos:

$$Ad^i = \frac{1}{\rho^i} A(u^{i+1} - u^i) = \frac{1}{\rho^i} (g_{i+1} - g_i), \quad i = 0, 1, \dots, n-1.$$

Poniendo $g_{i+1} = -d^{i+1} + \beta^i d^i$ y $g_i = -d^i + \beta^{i-1} d^{i-1}$, se deduce que Ad^i es una combinación lineal de d^{i-1} , d^i y d^{i+1} (para $i = 0$, $g_0 = -d^0$ y Ad^0 es una combinación lineal de d^1 y d^0).

Puesto que d^0, d^1, \dots, d^n son conjugadas se sabe que el punto u^{n+1} es el mínimo de J en la variedad $V_{n+1} = u^0 + \langle d^0, d^1, \dots, d^n \rangle$. Por tanto, $g_{n+1} = J'(u^{n+1})$ es ortogonal al subespacio $\langle d^0, d^1, \dots, d^n \rangle$. Como Ad^i pertenece al subespacio para $i = 0, 1, \dots, n-1$ se deduce $\langle Ad^i, g_{n+1} \rangle = 0$.

□

El algoritmo del *método iterativo de gradiente conjugado* para resolver sistemas de la forma $Au = b$ con matriz A simétrica y definida positiva es el siguiente:

(consideramos $r^n = -g_n = b - Au^n$)

1. Elegir $u_0 \in \mathbb{R}^N$, $\varepsilon, \delta > 0$.

Tomamos $r^0 = b - Au^0$, $d^0 = r^0$.

2. Dado $n \geq 0$, si $\|d^n\| < \delta$ paramos el algoritmo.

3. Calcular $\rho^n = \frac{\langle r^n, r^n \rangle}{\langle d^n, Ad^n \rangle}$, hacer $u^{n+1} = u^n + \rho^n d^n$ y $r^{n+1} = r^n - \rho^n Ad^n$.

Si $\|r^{n+1}\| < \varepsilon$ paramos el algoritmo.

4. En caso contrario calculamos $\beta^n = \frac{\langle r^{n+1}, r^{n+1} \rangle}{\langle r^n, r^n \rangle}$ y $d^{n+1} = r^{n+1} + \beta^n d^n$ y volvemos al paso 2.

Capítulo 5

Aplicación: resolución edp no lineal

Sea Ω un abierto acotado regular de \mathbb{R}^2 con $\partial\Omega = \Gamma_0 \cup \Gamma_1$ y $J : V \rightarrow \mathbb{R}$ el funcional dado por

$$J(v) = \frac{1}{2} \int_{\Omega} |\nabla v|^2 dx + \frac{1}{4} \int_{\Omega} |v|^4 dx + \frac{1}{2} \int_{\Omega} |v|^2 dx - \frac{1}{2} \int_{\Omega} f v dx \quad \forall v \in V,$$

definido sobre $V = \{v \in H^1(\Omega) : v|_{\Gamma_0} = 0\}$ espacio de Hilbert para el producto escalar

$$\langle v, w \rangle = \langle v, w \rangle_{H^1(\Omega)} = \langle v, w \rangle_{L^2(\Omega)} + \langle \nabla v, \nabla w \rangle_{L^2(\Omega)} \quad \forall v, w \in V \text{ y } f \in L^2(\Omega).$$

$$\|v\| = \|v\|_{H^1(\Omega)} = \|v\|_{1,2;\Omega} = \left(\int_{\Omega} |v(x)|^2 + |\nabla v(x)|^2 dx \right)^{1/2} \quad \forall v \in V.$$

Consideramos el problema de mínimos:

$$(P) \quad \text{Hallar } u \in V \text{ tal que } J(u) = \min_{v \in V} J(v).$$

Nuestro objetivo es resolver (P) mediante el Algoritmo del Gradiente con Paso Óptimo.

5.1. Marco teórico para la solución del problema

- En primer lugar comprobemos que nuestro problema (P) admite una única solución $u \in V$. Veamos que se cumplen las hipótesis del Teorema 2.17:

1. *J está bien definido.*

Es cierto pues cada una de las integrales está bien definida. Tomando $v \in V$, todas las integrales que intervienen en J son finitas, por lo que J está bien definido en V y toma valores en \mathbb{R} . Viendo la continuidad de J estamos comprobando lo que nos piden.

Aplicando la Desigualdad de Hölder (Proposición 1.30) tenemos:

$$\int_{\Omega} |\nabla v|^2 dx \leq \|\nabla v\|_{2,\Omega}^2 \leq \|v\|_{1,2,\Omega}^2,$$

$$\int_{\Omega} |v|^2 dx \leq \|v\|_{2,\Omega}^2 \leq \|v\|_{1,2,\Omega}^2,$$

$$\int_{\Omega} f v dx \leq \|f\|_{2,\Omega} \|v\|_{2,\Omega} \leq \|f\|_{2,\Omega} \|v\|_{1,2,\Omega}.$$

Y gracias a la Proposición 1.78 tenemos que $W^{1,2} \hookrightarrow L^4(\Omega)$, por lo que se cumple que

$$\int_{\Omega} |v^4| dx < \infty.$$

2. J es F -derivable en V .

Veamos en primer lugar que J es diferenciable en el sentido de Gâteaux, es decir, que existe el límite (3.3).

$$\begin{aligned} J(u + \delta w) &= \frac{1}{2} \int_{\Omega} |\nabla(u + \delta w)|^2 dx + \frac{1}{4} \int_{\Omega} |u + \delta w|^4 dx + \frac{1}{2} \int_{\Omega} |u + \delta w|^2 dx - \int_{\Omega} f(u + \delta w) \\ &= \frac{1}{2} \int_{\Omega} (|\nabla u|^2 + \delta^2 |\nabla w|^2 + 2\delta \nabla u \nabla w) dx + \frac{1}{4} \int_{\Omega} (u^4 + 6\delta^2 u^2 w^2 + 4\delta u^3 w) dx \\ &\quad + \frac{1}{4} \int_{\Omega} (4\delta^3 u w^3 + \delta^4 w^4) dx + \frac{1}{2} \int_{\Omega} (u^2 + \delta^2 w^2 + 2\delta u w) - \int_{\Omega} (f u + \delta f w) dx. \end{aligned}$$

$$\begin{aligned} \frac{J(u + \delta w) - J(u)}{\delta} &= \frac{1}{2} \int_{\Omega} (\delta |\nabla w|^2 + 2\nabla u \nabla w + \delta w^2 + 2u w) dx + \frac{1}{4} \int_{\Omega} (6\delta u^2 w^2 + 4u^3 w) dx \\ &\quad + \frac{1}{4} \int_{\Omega} (\delta^3 w^4 + 2\delta^2 u w^3) dx - \int_{\Omega} f w dx. \end{aligned}$$

Tomando límite cuando $\delta \rightarrow 0$ obtenemos:

$$\lim_{\delta \rightarrow 0} \frac{J(u + \delta w) - J(u)}{\delta} = \int_{\Omega} (\nabla u \nabla w + u w + u^3 w - f w) dx = \delta F(u),$$

por tanto, existe el límite y J es diferenciable en el sentido de Gâteaux.

Por la Proposición 3.8 sabemos que si J es diferenciable en el sentido de Fréchet entonces también lo es en sentido de Gâteaux. En tal caso, ambas derivadas coinciden. Consideramos la diferencial de J en el sentido de Gâteaux y vemos que es diferenciable Fréchet.

$$\begin{aligned} \frac{|o(w)|}{\|w\|} &= \frac{|J(u + w) - J(u) - \delta F(u)|}{\|w\|} \\ &= \frac{\left| \frac{1}{2} \int_{\Omega} |\nabla w|^2 dx + \frac{1}{2} \int_{\Omega} |w|^2 dx + \frac{1}{4} \int_{\Omega} (6u^2 w^2 + 2w^4 + 4u w^3) dx \right|}{\|w\|} \\ &\leq \frac{\frac{1}{2} \|\nabla w\|_{2,\Omega}^2 + \frac{1}{2} \|w\|_{2,\Omega}^2 + \frac{1}{2} \|w\|^4 + \|u\| \|w\|^3 + \frac{3}{2} \|u\|^2 \|w\|^2}{\|w\|} \\ &\leq \frac{\frac{1}{2} \|w\|^2 + \frac{1}{2} \|w\|^2 + \frac{1}{2} \|w\|^4 + \|u\| \|w\|^3 + \frac{3}{2} \|u\|^2 \|w\|^2}{\|w\|} \xrightarrow{w \rightarrow 0} 0 \end{aligned}$$

Hemos probado que J es diferenciable en el sentido de Fréchet y que

$$J'(u)(w) = \int_{\Omega} (\nabla u \nabla w + uw + u^3 w - fw) dx.$$

3. J es un funcional α -convexo en V (con $\alpha = 1$).

Por la Proposición 3.15 sabemos que si J es dos veces diferenciable de V en \mathbb{R} entonces que J sea α -convexo es equivalente a

$$J''(u)(w, w) \geq \alpha \|w\|^2 \quad \forall u, w \in V.$$

Basta comprobar que $J''(u)(w, w) \geq \|w\|^2$. Calculamos $J''(u)(w, w)$.

Primero la diferencial en el sentido de Gâteaux:

$$\frac{J'(u + \delta w_2) - J'(u)(w_1)}{\delta} = \int_{\Omega} (\nabla w_2 \nabla w_1 + w_1 w_2 + 3\delta u w_2^2 w_1 + 3u^2 w_2 w_1 + \delta w_2^2 w_1) dx \quad (5.1)$$

$$(5.1) \xrightarrow{\delta \rightarrow 0} \int_{\Omega} (\nabla w_2 \nabla w_1 + w_1 w_2 + 3u^2 w_2 w_1) dx = \delta^2 J.$$

Del mismo modo que para calcular J' , vemos que es diferenciable Fréchet.

$$\begin{aligned} \frac{|o(w)|}{\|w_2\|} &= \frac{|J'(u + w_2) - J'(u)(w_1) - \delta^2 J|}{\|w_2\|} = \frac{\left| \int_{\Omega} (3u w_2^2 w_1 + w_1 w_2^3) dx \right|}{\|w_2\|} \\ &\leq \frac{\int_{\Omega} (|3u w_2^2 w_1| + |w_1 w_2^3|) dx}{\|w_2\|} \end{aligned}$$

Aplicando la Desigualdas de Cauchy-Schwarz tenemos que

$$\frac{\int_{\Omega} (|3u w_2^2 w_1| + |w_1 w_2^3|) dx}{\|w_2\|} \xrightarrow{w_2 \rightarrow 0} 0$$

$$\text{Por tanto, } J''(u)(w_1, w_2) = \int_{\Omega} (\nabla w_2 \nabla w_1 + w_1 w_2 + 3u^2 w_2 w_1) dx.$$

$$J''(u)(w, w) = \int_{\Omega} (|\nabla w|^2 + w^2 + 3u^3 w^2) dx \geq \int_{\Omega} (|\nabla w|^2 + w^2) dx = \|w\|^2$$

Concluimos que J es un funcional α -convexo en V .

Como hemos probado que J es un funcional continuo α -convexo sobre V y considerando el espacio de Hilbert $H^1(\Omega)$, estamos en las condiciones para aplicar el Teorema 2.17. Por dicho teorema concluimos que existe un único mínimo u de J en V .

- Ahora observamos que el problema de mínimos (P) es equivalente al siguiente problema de contorno para una EDP no lineal:

$$(\tilde{P}) \quad \text{Hallar } u \in V \text{ tal que } \begin{cases} -\Delta u + u + u^3 & = f \text{ en } \Omega, \\ u & = 0 \text{ sobre } \Gamma_0, \\ \frac{\partial u}{\partial n} & = 0 \text{ sobre } \Gamma_1. \end{cases}$$

Multiplicamos la EDP de (\tilde{P}) por $v \in V$ e integramos en Ω .

$$\int_{\Omega} (-\Delta uv + uv + u^3v) dx = \int_{\Omega} fv dx. \quad (5.2)$$

Integrando por partes en (5.2) obtenemos

$$\int_{\Omega} (\nabla u \nabla v + uv + u^3v) dx = \int_{\Omega} fv dx \quad \forall v \in V,$$

pues $\int_{\partial\Omega} \frac{\partial u}{\partial n} v dS = \int_{\Gamma_0} \frac{\partial u}{\partial n} v dS + \int_{\Gamma_1} \frac{\partial u}{\partial n} v dS$ con $v|_{\Gamma_0} = 0$ y $\frac{\partial u}{\partial n} = 0$ en Γ_1 .

Por lo que consideramos

$$(FV) \quad \begin{cases} \text{Hallar } u \in V \text{ tal que} \\ \int_{\Omega} (\nabla u \nabla v + uv + u^3v) dx = \int_{\Omega} fv dx, \quad \forall v \in V. \end{cases}$$

Así tenemos $(\tilde{P}) \iff (FV)$.

Además, una condición necesaria para que $J(u) = \min_{v \in V} J(v)$ es $J'(u) = 0$. Previamente habíamos calculado

$$J'(u)(v) = \int_{\Omega} (\nabla u \nabla v + uv + u^3v - fv) dx.$$

$J'(u)(v) = 0 \iff (FV)$, concluimos por tanto la equivalencia entre (P) y (\tilde{P}) .

- Pasamos a aplicar el $(AGPO)$ visto en la Sección 4.1.

Veamos que estamos bajo las hipótesis del Teorema 4.1 de convergencia del $(AGPO)$. Para tener garantizada la convergencia del algoritmo falta probar que J' es globalmente Lipschitziana en los acotados de V . Hay que probar que:

$$\forall M > 0, \exists C_M > 0, \quad \|v\|, \|w\| \leq M \implies \|J'(v) - J'(w)\| \leq C_M \|v - w\|.$$

Aplicamos a $J' : V \rightarrow V'$ el Teorema del Valor Medio en un espacio de Banach y obtenemos lo siguiente para cierto ξ entre v y w :

$$\|J'(v) - J'(w)\|_{V'} \leq \|J''(\xi)\|_{\mathcal{B}(V)} \|v - w\|_V$$

donde usando la desigualdad de Hölder, las inyecciones de Sobolev y que v y w pertenecen a un acotado, se prueba que $\|J''(\xi)\| \leq C_M$ con $C_M \in \mathbb{R}$ y así concluimos que J' es globalmente Lipschitziana en acotados de V .

Para aplicar el algoritmo es necesario obtener $J'(u^n)$ y resolver el problema de minimización correspondiente en cada etapa.

Para el cálculo del gradiente, dado $w = u^n \in V$ queremos calcular $J'(w)$. Para ello identificamos V con V' .

Calcular $J'(w)$ es equivalente a hallar $g \in V$ tal que $\langle g, v \rangle_{H^1(\Omega)} = \langle J'(w), v \rangle_{V',V}$, $\forall v \in V$. A su vez, el problema anterior es equivalente a resolver el siguiente problema lineal:

$$(P_g) \quad \text{Hallar } g \in V \text{ tal que } \begin{cases} -\Delta g + g & = -\Delta w + w + w^3 - f & \text{en } \Omega, \\ g & = 0 & \text{sobre } \Gamma_0, \\ \frac{\partial g}{\partial n} & = \frac{\partial w}{\partial n} & \text{sobre } \Gamma_1. \end{cases}$$

donde el segundo miembro viene dado en la etapa anterior.

Veamos dicha equivalencia.

$$\langle g, v \rangle_{H^1(\Omega)} = \int_{\Omega} (gv + \nabla g \nabla v) dx.$$

$$J'(w)(v) = \langle J'(w), v \rangle_{V',V} = \int_{\Omega} (\nabla w \nabla v + vw + w^3 v - fv) dx.$$

Entonces,

$$\int_{\Omega} (gv + \nabla g \nabla v) dx = \int_{\Omega} (\nabla w \nabla v + vw + w^3 v - fv) dx. \quad (5.3)$$

Integrando por partes en el primer miembro de (5.3) obtenemos:

$$\int_{\Omega} (gv + \nabla g \nabla v) dx = \int_{\Omega} (gv - \Delta gv) dx + \int_{\Omega} \frac{\partial g}{\partial n} v dS = \int_{\Omega} (g - \Delta g)v dx + \int_{\Gamma_1} \frac{\partial g}{\partial n} v dS.$$

Integrando por partes en el segundo miembro de (5.3) obtenemos:

$$\begin{aligned} J'(w)(v) &= \int_{\Omega} (\nabla w \nabla v + vw + w^3 v - fv) dx = \int_{\Omega} (-\Delta w v + vw + w^3 v - fv) dx + \int_{\partial \Omega} \frac{\partial w}{\partial n} v dS \\ &= \int_{\Omega} (-\Delta w + w + w^3 - f)v dx + \int_{\Gamma_1} \frac{\partial w}{\partial n} v dS. \end{aligned}$$

Considerando $v \in H_0^1(\Omega) \subset V$, se tiene:

$$\int_{\Omega} (g - \Delta g)v dx = \int_{\Omega} (-\Delta w + w + w^3 - f)v dx$$

por lo que concluimos que

$$-\Delta g + g = -\Delta w + w + w^3 - f \quad \text{en } \Omega. \quad (5.4)$$

Ya hemos verificado que se cumple (5.4), veamos ahora las condiciones de contorno.

$$g \in V \implies g|_{\Gamma_0} = 0 \quad (g = 0 \text{ sobre } \Gamma_0).$$

Multiplicando (5.4) por $v \in V$ e integrando en Ω :

$$\int_{\Omega} (gv + \nabla g \nabla v) dx - \int_{\Gamma_1} \frac{\partial g}{\partial n} v dS = \int_{\Omega} (\nabla w \nabla v + wv + w^3 v - fv) dx - \int_{\Gamma_1} \frac{\partial w}{\partial n} v dS$$

de donde tiene que cumplirse $\int_{\Gamma_1} \frac{\partial g}{\partial n} v dS = \int_{\Gamma_1} \frac{\partial w}{\partial n} v dS \quad \forall v \in V$ para verificar (5.4).

Concluimos $\frac{\partial g}{\partial n} = \frac{\partial w}{\partial n}$ sobre Γ_1 .

Por otro lado, resolveremos el problema de minimización correspondiente a calcular ρ^n tal que $J(u^n - \rho^n J'(u^n)) = \inf_{\rho \in \mathbb{R}} J(u^n - \rho J'(u^n))$ en dos etapas. Una primera etapa de localización del mínimo mediante una Búsqueda no Restringida del intervalo con paso fijo y una segunda etapa de aproximación del mínimo utilizando el método de la Sección de Oro.

Primero debemos tener en cuenta observaciones y cálculos que serán necesarios para la localización del mínimo.

Planteamos el problema: obtener $\bar{\rho}$ tal que $J(u - \bar{\rho}g) = \min_{\rho \in \mathbb{R}} J(u - \rho g)$

Desarrollando la expresión $J(u - \rho g)$:

$$\begin{aligned} J(u - \rho g) &= \frac{1}{2} \int_{\Omega} |\nabla(u - \rho g)|^2 dx + \frac{1}{4} \int_{\Omega} |u - \rho g|^4 dx + \frac{1}{2} \int_{\Omega} |u - \rho g|^2 dx - \int_{\Omega} f(u - \rho g) dx \\ &= \frac{1}{2} \int_{\Omega} (|\nabla u|^2 + \rho^2 |\nabla g|^2 - 2\rho \nabla u \nabla g) dx + \frac{1}{2} \int_{\Omega} (u^2 + \rho^2 g^2 - 2\rho u g) dx \\ &\quad + \frac{1}{4} \int_{\Omega} (u^4 + 6\rho^2 u^2 g^2 - 4\rho u^3 g + \rho^4 g^4 - 4\rho^3 u g^3) dx - \int_{\Omega} f u dx + \int_{\Omega} \rho g f dx. \end{aligned}$$

Por tanto obtenemos el polinomio $p(\rho) = c_4 \rho^4 + c_3 \rho^3 + c_2 \rho^2 + c_1 \rho + c_0$, cuyos coeficientes son:

$$\begin{aligned} c_4 &= \frac{1}{4} \int_{\Omega} g^4 dx, \quad c_3 = - \int_{\Omega} u g^3 dx, \quad c_2 = \frac{1}{2} \int_{\Omega} |\nabla g|^2 dx + \frac{3}{2} \int_{\Omega} u^2 g^2 dx + \frac{1}{2} \int_{\Omega} g^2 dx, \\ c_1 &= - \int_{\Omega} \nabla u \nabla g dx - \int_{\Omega} u^3 g dx - \int_{\Omega} u g dx + \int_{\Omega} f g dx, \\ c_0 &= \frac{1}{2} \int_{\Omega} |\nabla u|^2 dx + \frac{1}{4} \int_{\Omega} u^4 dx + \frac{1}{2} \int_{\Omega} u^2 dx - \int_{\Omega} f u dx. \end{aligned}$$

Ahora estamos en condiciones de enunciar los dos algoritmos necesarios.

Algoritmo de Búsqueda no Restringida del Intervalo con Paso Fijo

Buscamos un intervalo $[a, b]$ tal que contenga x^* mínimo de la función $f(x)$.

1. Fijar un valor inicial $x_0 = 0$ y un paso s .
2. Calcular $f_0 = f(x_0)$.
3. Considerar $i = 1$ y $x_1 = x_0 + s$. Obtener $f_1 = f(x_1)$.
4. Si $f_1 = f_0$, entonces $x^* \in [x_0, x_1] = [a, b]$.
5. Si $f_1 < f_0$, se tiene que $x^* \geq x_0$. Tomar

$$x_{i+1} = x_i + s \quad \text{y} \quad f_{i+1} = f(x_{i+1}) \quad \text{para } i \geq 1,$$

hasta llegar a un i tal que $f_{i+1} \geq f_i$. Entonces $x^* \in [x_{i-1}, x_{i+1}] = [a, b]$.

6. Si $f_1 > f_0$, se tiene que $x^* \leq x_0$. Tomar $x_{-1} = x_0 - s$, $f_{-1} = f(x_{-1})$ y

$$x_{-(i+1)} = x_{-i} - s \quad \text{y} \quad f_{-(i+1)} = f(x_{-(i+1)}) \quad \text{para } i \geq 1,$$

hasta llegar a un i tal que $f_{-(i+1)} \geq f_{-i}$. Entonces $x^* \in [x_{-(i+1)}, x_{-(i-1)}] = [a, b]$.

Algoritmo del Método de la Sección de Oro

Suponemos conocido un intervalo $[a, b]$ donde se encuentra el mínimo.

1. Dados a y b tales que $a < b$, calcular $b - a$ y $\gamma = (1 + \sqrt{5})/2$.
2. Tomar $x_1 = a(\gamma - 1) + b(2 - \gamma)$ y $x_2 = a(2 - \gamma) + b(\gamma - 1)$.
3. Calcular $f_1 = f(x_1)$ y $f_2 = f(x_2)$.
4. Si $f_1 < f_2$, entonces $x^* \in [a, x_2]$. Tomar:

$$b = x_2, \quad x_2 = x_1, \quad x_1 = a(\gamma - 1) + b(2 - \gamma) \text{ y volver a (3).}$$

5. Si $f_1 = f_2$, entonces $x^* \in [x_1, x_2]$. Tomar:

$$a = x_1, \quad b = x_2, \quad x_1 = a(\gamma - 1) + b(2 - \gamma), \quad x_2 = a(2 - \gamma) + b(\gamma - 1) \text{ y volver a (3).}$$

6. Si $f_1 > f_2$, entonces $x^* \in [x_1, b]$. Tomar:

$$a = x_1, \quad x_1 = x_2, \quad x_2 = a(2 - \gamma) + b(\gamma - 1) \text{ y volver a (3).}$$

7. Repetir el proceso hasta llegar a un intervalo de amplitud menor o igual que un ε dado. Una aproximación del mínimo es, por ejemplo, el punto medio de dicho intervalo.

5.2. Resolución del problema con FreeFEM

5.2.1. En $\Omega = B(0, 1)$

Consideramos $\Omega = B(0, 1) \subset \mathbb{R}^2$ donde $\partial\Omega = \Gamma_0 \cup \Gamma_1$, con $\Gamma_1 = \{(x, y) \in \partial\Omega : x > 0\}$.

Empleando (AGPO):

Una vez tratadas todas las cuestiones teóricas anteriores, haremos un programa en FreeFEM para resolver el problema (P) utilizando el $(AGPO)$. Para hallar el mínimo del polinomio $p(\rho)$ discretizado, se realizará una primera fase de localización y posteriormente aproximación.

Los datos a considerar son:

- $n = 100$, el número de divisiones de $\partial\Omega$.
- $Nmax = 20$, el número máximo de iteraciones para el $(AGPO)$.
- $epsagpo = 10^{-3}$, la precisión para el test de parada de $(AGPO)$.
- Para el Algoritmo de Búsqueda no Restringida con Paso Fijo: $x_0 = 0$, $paso = 0.1$, $nmax = 10^6$.
- Para el Algoritmo de la Sección de Oro: $epsoro = 10^{-3}$ y $nmax = 10^6$.
- Para el $(AGPO)$: $u_0 = 0$.

El test de parada para el $(AGPO)$ es $\|J'(u_k)\|_V < epsagpo$, que está justificado por (4.4).

Las ejecuciones se harán para el segundo miembro f y la solución exacta uex en los siguientes casos:

$$\begin{cases} f &= 8(1 - 2x^2 - 2y^2) + (1 - x^2 - y^2)^2 + (1 - x^2 - y^2)^6, \\ uex &= (1 - x^2 - y^2)^2 \end{cases}$$

En las gráficas siguientes podemos observar tanto el dominio y su mallado, como las iteraciones del algoritmo. En este caso, son necesarias 3 iteraciones para que el error cometido sea menor que $epsagpo$.

En la primera iteración el algoritmo de búsqueda del intervalo con paso fijo nos da el intervalo $[0.8, 1]$ y el de la sección de oro nos dice que el mínimo se encuentra en $\rho_1 = 0.92717$ y $f(\rho_1) = -2.63333$.

En la segunda iteración el intervalo que contiene al mínimo es $[0.8, 1]$ y este se encuentra en $\rho_2 = 0.897369$ y $f(\rho_2) = -2.63832$.

Y por último, en la tercera iteración el intervalo que contiene al mínimo es $[0.7, 0.9]$ y este se encuentra en $\rho_3 = 0.824301$ y $f(\rho_3) = -2.6384$.

Apreciando la escala de colores, podemos observar donde se encuentran los valores más bajos (colores naranjas) y los más altos (colores rojos y rosas).

Frontera del dominio

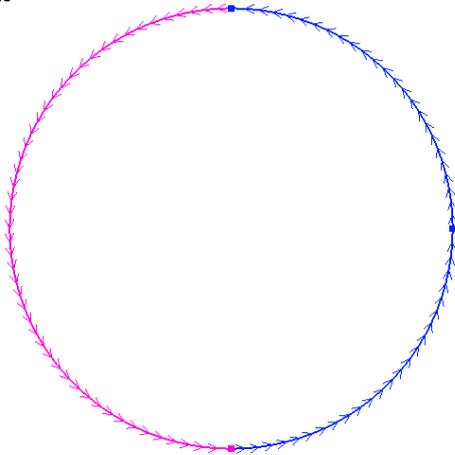


Figura 5.1: Frontera del dominio

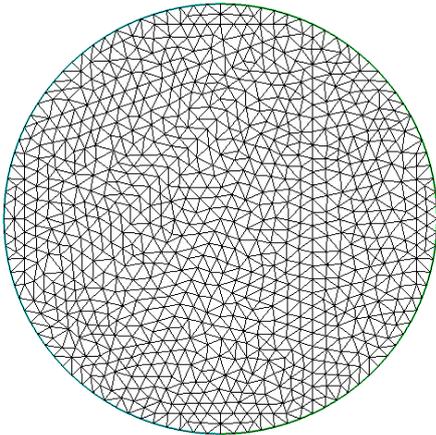


Figura 5.2: Malla del dominio

f: funcion segundo miembro

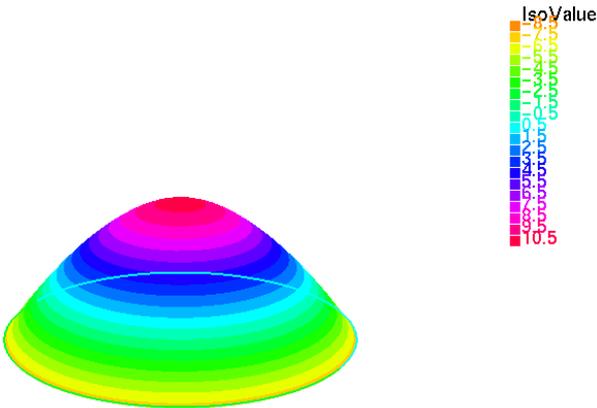


Figura 5.3: Función f

u_0 : u inicial

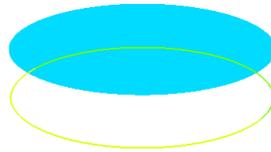


Figura 5.4: Función u_0 : inicialización AGPO

Iteracion k= 1, error=5.48079

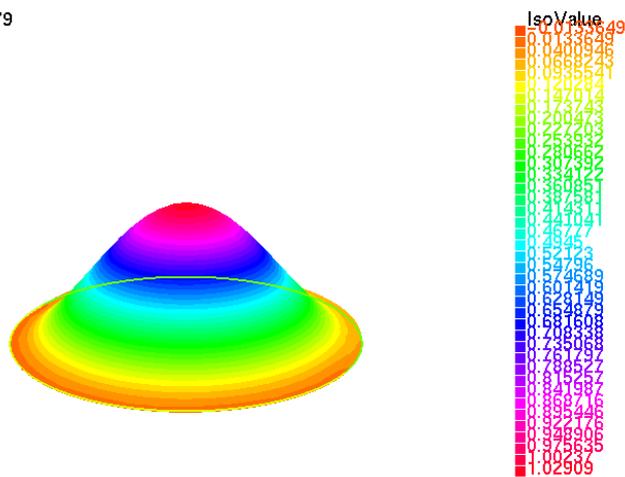


Figura 5.5: Primera iteracion del AGPO

Iteracion k= 2, error=0.0111401

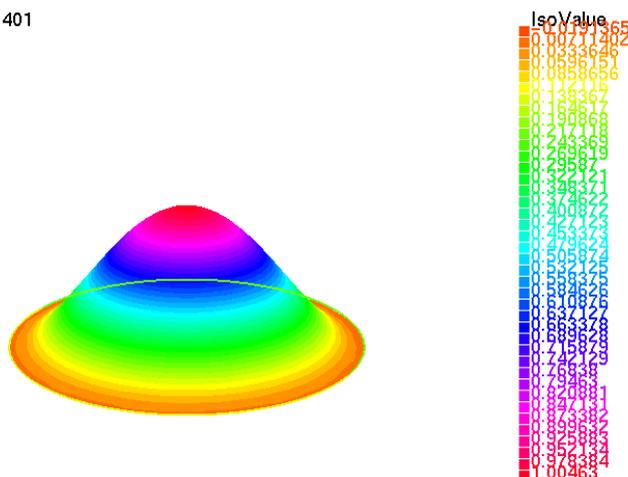


Figura 5.6: Segunda iteracion del AGPO

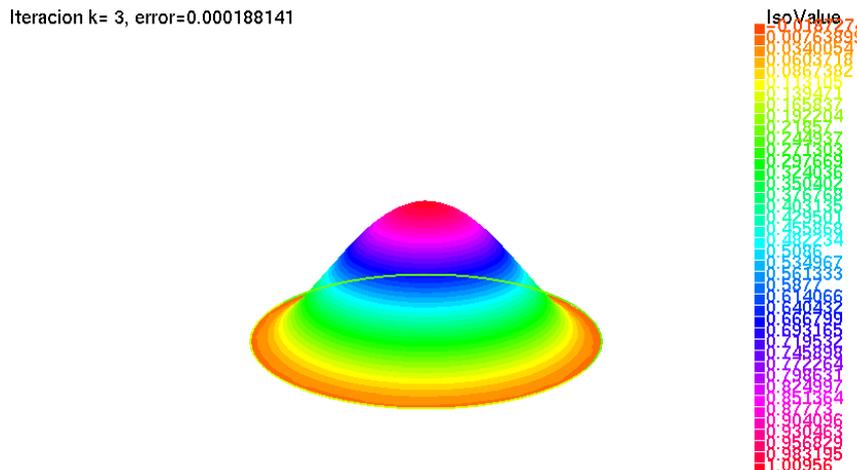


Figura 5.7: Tercera iteracion del AGPO

Además, $\|uex\|_{L^2(\Omega)} = 0.796376$, $\|u^3\|_{L^2(\Omega)} = 0.782665$ y $\|u^3 - uex\|_{L^2(\Omega)} = 0.00694$, por lo que verificamos la convergencia del algoritmo.

Si tomamos, por ejemplo, $n = 140$, calculamos $\|u^3 - uex\|_{L^2(\Omega)} = 0.00298$ y comprobamos que dicho error va disminuyendo respecto al caso anterior ($n = 100$) al considerar una malla más fina. Para este caso, también han sido necesarias 3 iteraciones.

Empleando Paso Fijo:

Queremos resolver nuestro problema con el algoritmo de gradiente con paso fijo (visto en la Sección 4.2) en lugar de con (AGPO) para comparar ambos métodos. Tomamos $Nmax = 50$.

Con este algoritmo son necesarias 38 iteraciones, muchas más que en el caso anterior y además $\|u^{38} - uex\|_{L^2(\Omega)} = 0.01128$, por lo que la aproximación era mejor con el (AGPO).

Veamos algunas de las iteraciones:

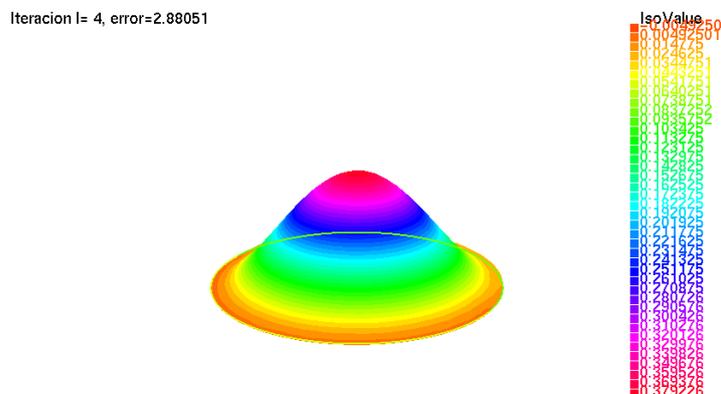


Figura 5.8: 4ª iteración del Algoritmo de gradiente con paso fijo

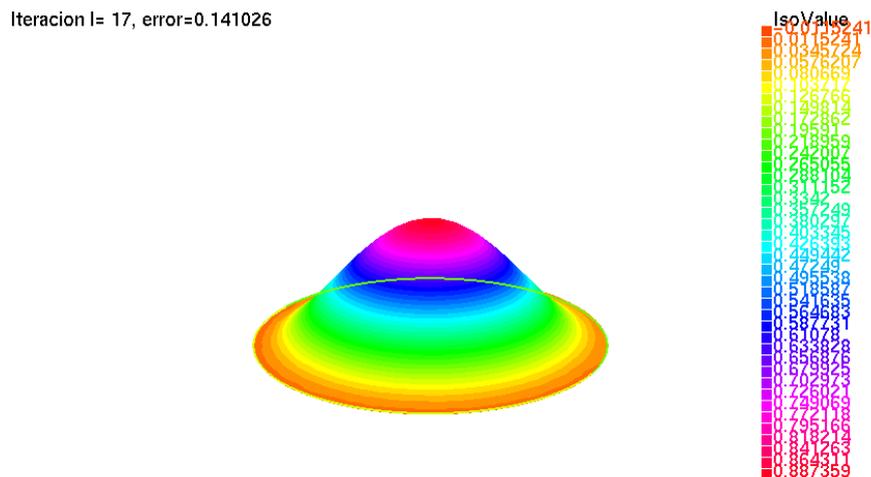


Figura 5.9: 17ª iteración del Algoritmo de gradiente con paso fijo

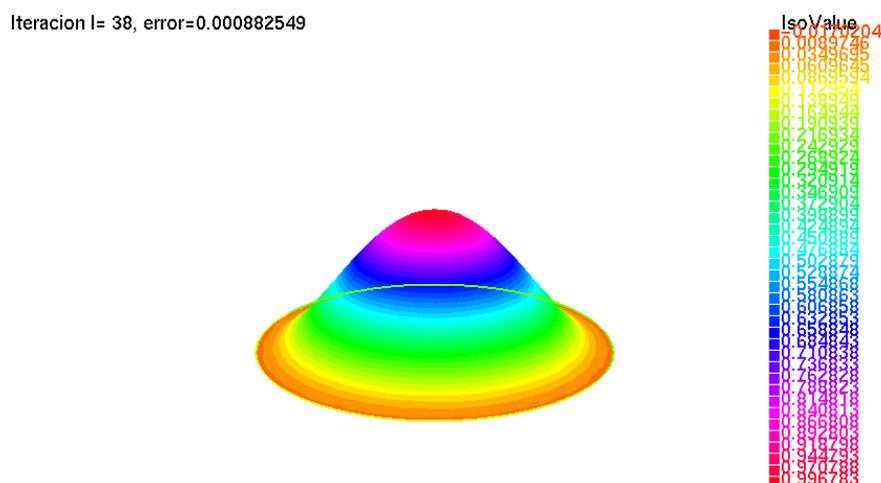


Figura 5.10: 38ª iteración del Algoritmo de gradiente con paso fijo

5.2.2. En Ω una elipse con un agujero

También podemos aplicar el código de FreeFEM de la resolución mediante el (AGPO) a una función f general, de la cual no conocemos la solución. Además, consideramos el siguiente dominio: $\Omega = \Omega_1 \setminus B(0, 0.25) \subset \mathbb{R}^2$ donde $\Omega_1 = \{(x, y) \in \mathbb{R}^2 \mid x = \cos(t), y = 0.5 \sin(t)\}$, $\partial\Omega = \Gamma_0 \cup \Gamma_1$ con $\Gamma_1 = \partial B(0, 0.25)$.

Volvemos a considerar el funcional J definido al comienzo del capítulo y el espacio de Hilbert $V = \{v \in H^1(\Omega) : v|_{\Gamma_0} = 0\}$ dotado del producto escalar $\langle v, w \rangle = \langle v, w \rangle_{H^1(\Omega)} \quad \forall v, w \in V$.

Sea el problema de mínimos:

$$(P) \quad \text{Hallar } u \in V \text{ tal que } J(u) = \min_{v \in V} J(v).$$

Nuestro objetivo es resolver (P) mediante el Algoritmo del Gradiente con Paso Óptimo.

En este caso, tomamos por ejemplo $f_1 = 10(x^2 + x + y^2)$. Todos los argumentos probados en la Sección 5.1 también son válidos para este dominio y esta función f_1 . Además, consideramos $\epsilon_{\text{stop}} = 10^{-6}$.

En las gráficas siguientes podemos ver tanto el dominio y su mallado, como las iteraciones del algoritmo. Serán necesarias también 3 iteraciones.

Frontera del dominio

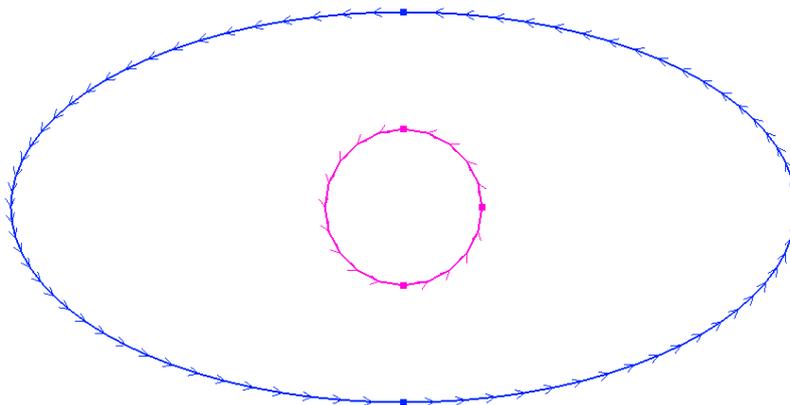


Figura 5.11: Frontera elipse

Malla del dominio

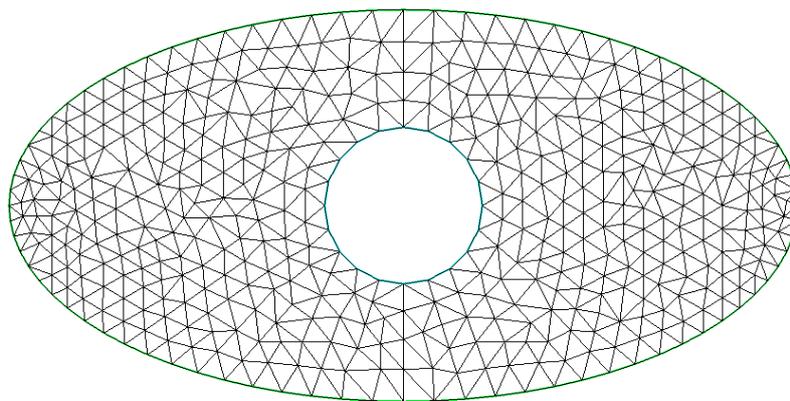


Figura 5.12: Malla de la elipse

En la primera iteración el algoritmo de búsqueda del intervalo con paso fijo nos da el intervalo $[0.9, 1.1]$ y el de la sección de oro nos dice que el mínimo se encuentra en $\rho_1 = 0.995359$ y $f(\rho_1) = -0.699045$.

En la segunda iteración el intervalo que contiene al mínimo es $[0.9, 1.1]$ y este se encuentra en $\rho_2 = 0.993496$ y $f(\rho_2) = -0.699048$.

En la tercera iteración el intervalo que contiene al mínimo es $[0.9, 1.1]$ y este se encuentra en $\rho_2 = 0.987347$ y $f(\rho_2) = -0.699048$.

Finalmente, $\|u^3\|_{L^2(\Omega)} = 0.231649$.

f: funcion segundo miembro

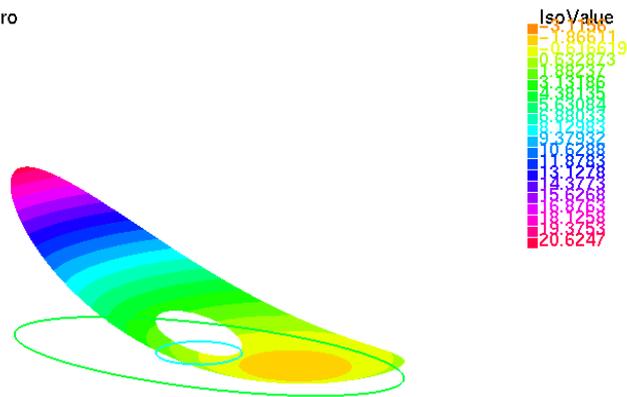


Figura 5.13: Función f_1

u_0 : u inicial

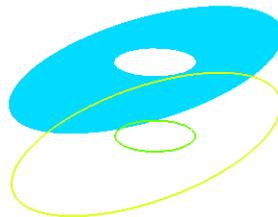


Figura 5.14: Función u_0 : inicialización AGPO (para la elipse y f_1)

Iteracion k= 1, error=1.40141

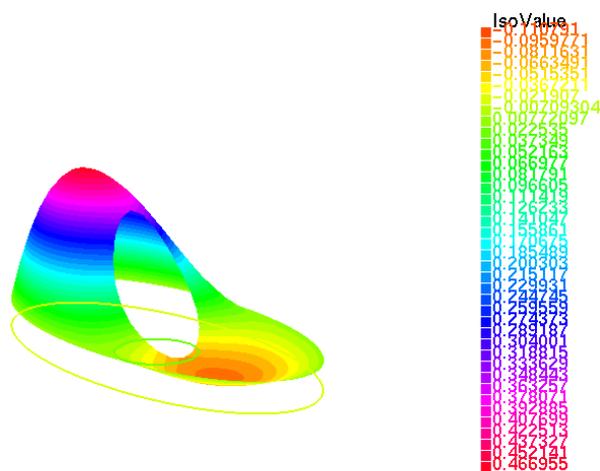


Figura 5.15: Primera iteracion del AGPO (para la elipse y f_1)

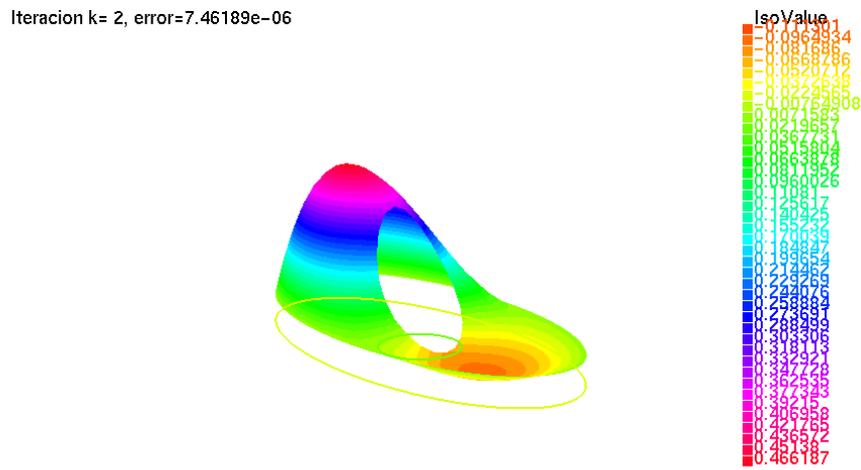


Figura 5.16: Segunda iteracion del AGPO (para la elipse y f_1)

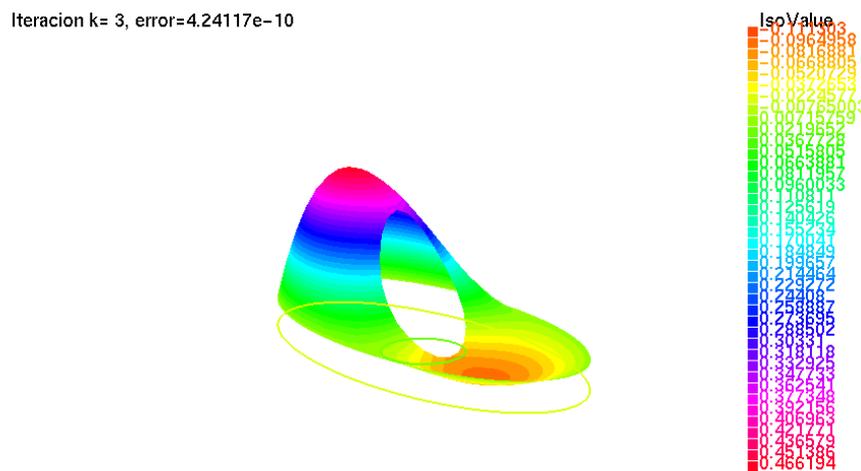


Figura 5.17: Tercera iteracion del AGPO (para la elipse y f_1)

Anexo

Adjuntamos a continuación los códigos de FreeFEM empleados para la resolución de los distintos problemas del Capítulo 5. Comenzamos con el programa utilizado para resolver el problema original planteado en $B(0,1)$ y la función f .

```
1 // RESOLUCION DEL PROBLEMA PLANTEADO.
2 //DOMINIO: B(0,1)
3
4 //DATOS
5
6 real R=1.; //Radio de la bola
7 real pi2 = 0.5*pi; // pi/2
8 int n =100; //Numero de divisiones en la frontera de Omega
9 int L0=9, L1=12; //Etiquetas para la frontera
10
11 //DATOS PARA LA BUSQUEDA NO RESTRINGIDA
12 real a,b; // Extremos del intervalo buscado
13 real x0=0.; // Valor inicial
14 real s=0.1; // Paso
15 real nmax= 1.e6; // Numero maximo de iteraciones
16
17
18 //DATOS PARA LA SECCION DE ORO
19 real gamma=0.5*(1+sqrt(5)); // Constante gamma
20 real gamma1= gamma -1, gamma2=2-gamma;
21 real x1= a*gamma1 +b*gamma2;
22 real x2= a*gamma2 +b*gamma1;
23 real epsoro= 1.e-3; // Epsilon: precision del test de parada, para medir la
    amplitud del intervalo
24
25 //DATOS (AGPO)
26 func u0 = 0; // Valor inicial
27 real Nmax =20; // Numero maximo de iteraciones
28 int k; // Parametro para las iteraciones
29 real epsagpo= 1.e-3; // Epsilon: precision del test de parada
30
31
32 //Frontera del dominio Omega=B(0,1)
```

```

33 border C0 (t=0,pi2) {x=R*cos(t); y=R*sin(t) ; label=L0;};
34 border Gamma1 (t=pi2,3*pi2) {x=R*cos(t); y=R*sin(t) ; label=L1;};
35 border C2 (t=3*pi2,4*pi2) {x=R*cos(t); y=R*sin(t) ; label=L0;};
36
37 // Dibujamos la frontera del dominio
38 plot(C0(0.25*n)+Gamma1(0.5*n)+C2(0.25*n),cmm = " Frontera del dominio ", wait=
    true);
39
40
41 //Mallado del dominio
42 mesh Th = buildmesh( C0(0.25*n)+Gamma1(0.5*n)+C2(0.25*n));
43 plot(Th, cmm = " Malla del dominio ", wait = true);
44
45
46 //funciones
47 func uex = (1-x*x-y*y)*(1-x*x-y*y); // Solucion exacta
48 func f = 8*(1-2*x*x-2*y*y) + uex + uex*uex*uex; // Funcion del segundo miembro
49
50
51 //Espacio de elementos finitos
52 fespace Vh(Th, P1);
53 Vh g, v; // Variables asociadas a la formulacion variacional del problema
    Gradiente
54 Vh u; //
55 Vh uexh = uex; // Funcion uex como elemento de Vh
56 Vh fh = f; // Funcion f como elemento de Vh
57
58
59 //Dibujamos la funcion f como elemento de Vh
60 plot(fh, value = true, fill = true, cmm = "f: funcion segundo miembro",
    = true);
61
62
63 //CALCULO DEL GRADIENTE
64 // Tenemos que resolver el siguiente problema. Para ello definimos previamente
    el macro grad:
65
66 macro grad(u) [dx(u),dy(u)] // EOM
67
68 problem Gradiente (g,v,init = 1) =
69     int2d(Th)(grad(g)* grad(v)) +int2d(Th)(g*v)
70     - int2d(Th)(grad(u)*grad(v))
71     -int2d(Th)(u*v) -int2d(Th)(u*u*u*v) +int2d(Th)(fh*v)
72     + on(L0,g=0.);
73
74
75 //FUNCION AUXILIAR PARA EL CALCULO DE LOS COEFICIENTES

```

```

76 // Con esta funcion coefpol calculamos los 5 coeficientes necesarios para el
    polinomio
77 // el cual queremos calcular su minimo :
78 //p(rho)= c4*x^4 + c3*x^3 + c2*x^2 + c1*x + c0,
79 func real[int] coefpol(real[int] vu, real[int] vg)
80 {
81   Vh u, g;
82   u[]=vu;
83   g[]=vg;
84
85   real c0,c1,c2,c3,c4;
86   // Definimos los 5 coeficientes
87   c4= 0.25*int2d(Th)(g*g*g*g);
88   c3= - int2d(Th)(u*g*g*g);
89   c2= 0.5* int2d(Th)(grad(g)*grad(g))
90     +1.5*int2d(Th)(u*u*g*g) + 0.5*int2d(Th)(g*g);
91
92   c1= -int2d(Th)(grad(g)*grad(u))
93     -int2d(Th)(u*u*u*g)-int2d(Th)(u*g)+int2d(Th)(fh*g);
94
95   c0= 0.5* int2d(Th)(grad(u)*grad(u))
96     + 0.25*int2d(Th)(u*u*u*u) + 0.5*int2d(Th)(u*u)-int2d(Th)(fh*u);
97
98   real[int] b=[c4,c3,c2,c1,c0]; // Devuelve un vector
99   return b;
100 }
101
102
103 // Resolvemos finalmente mediante el Algoritmo de Gradiente con Paso Optimo.
104
105 // (AGPO)
106 u=u0; // Inicializamos
107 // Pintamos la solucion inicial
108 plot(u, nbiso = 40, dim = 3, fill=true,
109       cmm = " u0 : u inicial", wait = true);
110
111 //Comenzamos el bucle con las iteraciones:
112 for ( k=1; k<Nmax+1; k++){
113
114   Gradiente; // Resolvemos el problema Gradiente y obtenemos g.
115
116   //Algoritmo de Busqueda no Restringida del intervalo con paso fijo
117   // Definicion de la funcion f(x) que queremos minimizar, en nuestro caso es el
    polinomio p(rho):
118   func real f(real x) {
119     return coefpol(u[],g[])[0]*x^4 + coefpol(u[],g[])[1]*x^3 + coefpol(u[],g[])[
    2]*x^2 + coefpol(u[],g[])[3]*x + coefpol(u[],g[])[4];
120   }

```

```
121
122 // Calculo del minimo de la funcion
123 real f0 = f(x0);
124 real f1 = f(x0 + s);
125
126 if (f1 == f0) {
127     // El minimo esta en [x0, x0 + s]
128     a = x0;
129     b = x0 + s;
130 }else if (f1 < f0) {
131     real xi = x0 + s;
132     real fi = f1;
133     real xiprev = x0;
134     real fiprev = f0;
135     int i;
136     while ( i<nmax+1 && fi < fiprev ) {
137         xiprev = x0 + (i-1) * s;;
138         fiprev = f(xiprev);
139         xi = x0 + i * s;
140         fi = f(xi);
141         i++;
142     }
143     // El minimo esta en [x{i-1}, x{i+1}]
144     a = xiprev -s;
145     b = xi;
146 } else {
147     real xi = x0 - s;
148     real fi = f(xi);
149     real xiprev = x0;
150     real fiprev = f0;
151     int i = 1;
152
153     while (i<nmax+1 && fi < fiprev) {
154         xiprev = x0 - (i-1) * s;
155         fiprev = f(xiprev);
156         xi = x0 - i * s;
157         fi = f(xi);
158         i++;
159     }
160
161     // El minimo esta en [x{-(i+1)}, x{-(i-1)}]
162     a = xi ;
163     b = xiprev + s;
164 }
165
166 // Imprimir el resultado
167 cout << "Intervalo que contiene el minimo: [" << a << ", " << b << "]" << endl;
168
```

```

169
170 // Algoritmo de la seccion de oro
171 func real SeccionOro(real ad, real bd, real epsorod) {
172     real a=ad, b=bd, epsoro=epsorod;
173     real gamma1= gamma -1, gamma2=2-gamma;
174     real x1= a*gamma1 +b*gamma2;
175     real x2= a*gamma2 +b*gamma1;
176     real fx1 = f(x1);
177     real fx2 = f(x2);
178
179     while (abs(b-a) > epsoro){
180         if (fx1==fx2){
181             a=x1;
182             b=x2;
183             x1= a*gamma1 +b*gamma2;
184             x2= a*gamma2 +b*gamma1;
185         }
186         if (fx1 < fx2 ) {
187             b = x2;
188             x2 = x1;
189             x1= a*gamma1 +b*gamma2;
190             fx2 = fx1;
191             fx1 = f(x1);
192         } else {
193             a = x1;
194             x1 = x2;
195             x2= a*gamma2 +b*gamma1;
196             fx1 = fx2;
197             fx2 = f(x2);
198         }
199     }
200     real xm=(b + a)*0.5;
201     return xm;
202 }
203
204 real rhok = SeccionOro(a, b, epsoro); // Asi tenemos calculado el paso rho_k
205 // Imprimir resultado
206 cout << "Minimo encontrado en x = " << rhok << endl;
207 real fx=f(rhok);
208 cout << "Y vale f(x)= " << fx << endl;
209
210 u=u - rhok*g; // Calculamos la nueva iteracion con el (AGPO)
211 real error;
212 error = int2d(Th)(g*g + grad(g)'*grad(g)); // Definimos el error
213 plot(u, nbiso = 40, dim = 3, cmm="Iteracion k= "+ k+", error="+error, fill=
214     true, value = true,
215     wait=false); //Dibujamos la solucion en cada iteracion

```

```

215   if (error<epsagpo){
216       break;} //Test de parada
217
218   }
219
220   real error1;
221   error1=sqrt(int2d(Th)(u^2));
222   real error2;
223   error2=sqrt(int2d(Th)(uex^2));
224   cout << "Norma en L^2 de la solucion calculada " << error1 << endl;
225   cout << "Norma en L^2 de la solucion exacta " << error2 << endl;
226   real error3;
227   error3=sqrt(int2d(Th)((u-uex)^2));
228   cout << "Norma en L^2 de u-uex " << error3 << endl;
229
230

```

Vemos a continuación el código empleado para resolver el mismo problema pero mediante el algoritmo de paso fijo.

```

1 // RESOLUCION DEL PROBLEMA USANDO PASO FIJO.
2 //DOMINIO: B(0,1)
3
4 //DATOS
5
6 real R=1.; //Radio de la bola
7 real pi2 = 0.5*pi; // pi/2
8 int n =100; //numero de divisiones en la frontera de Omega
9 int L0=9, L1=12; //Etiquetas para la frontera
10
11
12 //DATOS GRADIENTE CON PASO FIJO
13 func u0 = 0; // Valor inicial
14 real Nmax =50; // Numero maximo de iteracione
15 int l; // Parametro para las iteraciones
16 real epsagpo= 1.e-3; // Epsilon: precision del test de parada
17 real rho=0.1; // Paso (fijo)
18
19
20 //Frontera del dominio Omega=B(0,1)
21 border C0 (t=0,pi2) {x=R*cos(t); y=R*sin(t) ; label=L0;};
22 border Gamma1 (t=pi2,3*pi2) {x=R*cos(t); y=R*sin(t) ; label=L1;};
23 border C2 (t=3*pi2,4*pi2) {x=R*cos(t); y=R*sin(t) ; label=L0;};
24
25 plot(C0(0.25*n)+Gamma1(0.5*n)+C2(0.25*n), wait=true);
26
27

```

```

28 //Mallado del dominio
29 mesh Th = buildmesh( C0(0.25*n)+Gamma1(0.5*n)+C2(0.25*n));
30 plot(Th, cmm = " Malla inicial ", wait = true);
31
32
33 //funciones
34 func uex = (1-x*x-y*y)*(1-x*x-y*y); //solucion exacta
35 func f = 8*(1-2*x*x-2*y*y) + uex + uex*uex*uex;
36
37
38 //Espacio de elementos finitos
39 fespace Vh(Th, P1);
40 Vh g, v; //Variables asociadas a la formulacion variacional
41 Vh u;
42 Vh uexh = uex; // Funcion uex como elemento de Vh
43 Vh fh = f; // Funcion f como elemento de Vh
44
45 //Dibujamos la funcion f como elemento de Vh
46 plot(fh, value = true, fill = true, cmm = "f", wait = true);
47
48
49 //CALCULO DEL GRADIENTE
50 // Tenemos que resolver el siguiente problema. Para ello definimos previamente
    el macro grad:
51
52 macro grad(u) [dx(u),dy(u)]// EOM
53
54 problem Gradiente (g,v,init = 1) =
55     int2d(Th)(grad(g)*grad(v)) +int2d(Th)(g*v)
56     - int2d(Th)(grad(u)*grad(v))
57     -int2d(Th)(u*v) -int2d(Th)(u*u*u*v) +int2d(Th)(fh*v)
58     + on(L0,g=0.);
59
60
61 // ALGORITMO DE GRADIENTE CON PASO FIJO
62 u=u0;// Inicializamos
63 // Pintamos la solucion inicial
64 plot(u, nbiso = 40, dim = 3, fill=true,
65     cmm = " u0 : u inicial", wait = true);
66
67 //Comenzamos el bucle con las iteraciones:
68
69 for ( l=1; l<Nmax+1; l++){
70     Gradiente; // Resolvemos el problema Gradiente y obtenemos g.
71     u=u - rho*g; // Definimos la nueva iteracion
72     real error;
73     error = int2d(Th)(g*g + grad(g)*grad(g));
74     plot(u, nbiso = 40, dim = 3, cmm="Iteracion l= "+ l+", error="+error, fill=

```

```

    true, value = true,
75         wait=false); //Dibujamos la solucion en cada iteracion
76 if (error<epsagpo){
77     break;} //Test de parada
78
79 }
80
81 real error1;
82 error1=sqrt(int2d(Th)(u^2));
83 real error2;
84 error2=sqrt(int2d(Th)(uex^2));
85 cout << "Norma en L^2 de la solucion calculada " << error1 << endl;
86 cout << "Norma en L^2 de la solucion exacta " << error2 << endl;
87 real error3;
88 error3=sqrt(int2d(Th)((u-uex)^2));
89 cout << "Norma en L^2 de u-uex " << error3 << endl;

```

Finalmente, el código empleado para resolver el problema planteado en la elipse con un agujero y la función f_1 , mediante (AGPO), es el mismo que el primero salvo en la definición del dominio y de la triangulación.

```

1 //Frontera del dominio Omega
2 //Elipse
3 border C0 (t=0,pi2) {x=cos(t); y=0.5*sin(t) ; label=L0;};
4 border Gamma1 (t=pi2,3*pi2) {x=cos(t); y=0.5*sin(t) ; label=L0;};
5 border C2 (t=3*pi2,4*pi2) {x=cos(t); y=0.5*sin(t) ; label=L0;};
6 //Bola
7 border C3 (t=0,pi2) {x=0.2*cos(t); y=0.2*sin(t) ; label=L1;};
8 border C4 (t=pi2,3*pi2) {x=0.2*cos(t); y=0.2*sin(t) ; label=L1;};
9 border C5 (t=3*pi2,4*pi2) {x=0.2*cos(t); y=0.2*sin(t) ; label=L1;};
10 // Dibujamos la frontera del dominio
11 plot(C0(0.2*n)+Gamma1(0.4*n)+C2(0.2*n)+C3(-0.05*n)+C4(-0.1*n)+C5(-0.05*n),cmm =
    " Frontera del dominio ", wait=true);
12
13
14 //Mallado del dominio
15 mesh Th = buildmesh( C0(0.2*n)+Gamma1(0.4*n)+C2(0.2*n)+C3(-0.05*n)+C4(-0.1*n)+C5
    (-0.05*n));

```

Bibliografía

- [1] G.N. Gatica, *Introducción al Análisis Funcional. Teoría y Aplicaciones*. Reverté, 2014.
- [2] H. Brezis, *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Springer, 2010.
- [3] V. Serov, *Fourier Series, Fourier Transform and Their Applications to Mathematical Physics*. Springer, 2017.
- [4] W. Rudin, *Real and Complex Analysis*, 3^a ed. New York: McGraw-Hill, 1987.
- [5] J.-M. Bony, *Cours d'analyse : Théorie des distributions et analyse de Fourier*. Ellipses Marketing, 2000.
- [6] H. Brezis, *Analyse fonctionnelle: Théorie et applications*. Paris: Masson, 1983.
- [7] G. Allaire, *Numerical Analysis and Optimization: An Introduction to Mathematical Modeling and Numerical Simulation*. Oxford University Press, USA, 2007.
- [8] E. Kreyszig, *Introductory Functional Analysis with Applications*. John Wiley and Sons Ltd, 1978.
- [9] M.L. Stockdale, *El problema del viajante: un algoritmo heurístico y una aplicación*. Tesis de Licenciatura, 2011.
- [10] J.-L. Lions, *Contrôle Optimal de Systèmes Gouvernés par des Equations aux Dérivées Partielles*. París: Dunod, 1968.
- [11] I. Ekeland and R. Temam, *Convex Analysis and Variational Problems*. Philadelphia: Society for Industrial and Applied Mathematics, 1999.
- [12] J. L. Troutman, *Variational Calculus and Optimal Control: Optimization With Elementary Convexity*, 2^a ed. New York: Springer, 1996.
- [13] J. M. Viaño, *Lecciones de métodos numéricos 4.: Optimización*. Andavira Editora, 1995.