

UNIVERSIDAD NACIONAL DEL LITORAL
FACULTAD DE INGENIERÍA QUÍMICA

TESIS PRESENTADA COMO PARTE DE LOS REQUISITOS DE LA UNIVERSIDAD NACIONAL DEL
LITORAL PARA LA OBTENCIÓN DEL GRADO ACADÉMICO DE

Magíster en Matemática

EN EL CAMPO DE: **Matemática Aplicada**

TÍTULO DE LA TESIS:

**Regularización de problemas inversos mal condicionados
mediante la minimización de funcionales de tipo
Tikhonov-Phillips doblemente generalizados**

INSTITUCIÓN DONDE SE REALIZÓ:

Instituto de Matemática Aplicada del Litoral (CONICET-UNL)
Facultad de Bioquímica y Ciencias Biológicas (UNL)

AUTOR:

Lic. María Josefina Carrió

DIRECTOR DE TESIS:

Dra. Gisela Luciana Mazzieri

CODIRECTOR DE TESIS:

Dra. Karina Guadalupe Temperini

MIEMBROS DEL JURADO:

Dra. Andrea Bergesio

Dra. Marcela Morvidone

Dr. Sebastian Pauletti

AÑO DE PRESENTACIÓN: 2019

A mamá y papá...

Agradecimientos

*“Un hilo rojo invisible conecta a aquellos
que están destinados a encontrarse,
sin importar tiempo, lugar o circunstancias.*

*El hilo se puede estirar o contraer,
pero nunca romper...”*

Mitología china

Gracias a todos los que hicieron que esto sea posible. GRACIAS!!!

Índice general

<i>Agradecimientos</i>	I
Resumen	V
Introducción	VII
Capítulo 1. Problemas inversos desde el enfoque determinístico	1
1.1. Teoría Clásica	3
1.2. Regularización de tipo Tikhonov-Phillips generalizado	10
1.3. Regularización de tipo Tikhonov-Phillips doblemente generalizado	16
Capítulo 2. Problemas inversos desde el enfoque estocástico	29
2.1. Tareas	31
2.1.1. Densidad a-priori	31
2.1.2. Función de verosimilitud	36
2.1.3. Densidad a-posteriori	37
2.2. Vínculo entre los enfoques determinístico y estocástico	38
2.3. Una regla de elección del parámetro heurística	43
2.4. Aplicaciones	49
2.4.1. Señales	50
2.4.2. Imágenes	54
Capítulo 3. Estimaciones del error a-posteriori	61
3.1. Regularización de tipo Tikhonov-Phillips generalizado- p	62
3.2. Regularización de tipo Tikhonov-Phillips doblemente generalizado- p	67
Conclusiones y trabajos futuros	79
Bibliografía	83

Resumen

En este trabajo de tesis se aborda el estudio de la regularización de problemas inversos mal condicionados desde el enfoque determinístico de su resolución y, en este contexto, se presentan los métodos de Tikhonov-Phillips en sus versiones clásica, generalizada (es decir, con término de fidelidad cuadrático y penalizante generalizado) y doblemente generalizada (es decir, con término de fidelidad y penalizante generalizados). Se muestra que estos métodos constituyen regularizaciones en espacios de Banach. Para ello, para el caso doblemente generalizado, se prueban la existencia y unicidad de soluciones regularizadas, la existencia de solución de mínimo penalizante y, finalmente, la convergencia de aquellas a esta. El resultado de convergencia se obtiene para el caso en que el parámetro de regularización es elegido mediante una regla a-priori.

Por otra parte, se presenta el enfoque estadístico para la resolución de los problemas inversos y su principal vínculo con el enfoque determinístico clásico. Bajo un modelo lineal con ruido aditivo, a partir de diferentes ejemplos, se muestra cómo la densidad de la variable aleatoria que representa el error se relaciona con el término de fidelidad de un funcional de tipo Tikhonov-Phillips doblemente generalizado y cómo la densidad a-priori de la variable incógnita está estrechamente relacionada con el penalizante de dicho funcional. Más precisamente, se prueba que bajo el supuesto de distribuciones a-priori (para el ruido y la variable de interés) pertenecientes a familias exponenciales, el estimador máximo a-posteriori calculado desde el enfoque estocástico coincide con la solución regularizada obtenida mediante la minimización de un funcional de tipo Tikhonov-Phillips doblemente generalizado.

A su vez, se estudia una regla de elección del parámetro heurística presentada por K. Ito, B. Jin y J. Zou en 2011 cuya construcción se basa en el enfoque estocástico de la resolución de un problema inverso, más precisamente, en los modelos jerárquicos. Además, se presentan algunas aplicaciones a problemas inversos en restauración de señales e imágenes donde se utiliza dicha regla para determinar el parámetro óptimo en cada caso. Los resultados numéricos obtenidos permiten una mejor visualización de los resultados teóricos.

Cuando la regla de elección del parámetro es heurística, no podemos esperar obtener resultados de convergencia. En este caso, se prueban estimaciones del error a-posteriori cuando

la solución regularizada resulta de la minimización de funcionales de tipo Tikhonov-Phillips generalizado o doblemente generalizado. Esta estimación se particulariza a la regla heurística dada en 2011 por K. Ito, B. Jin y J. Zou dejando en evidencia que resulta mejor que la dada por los autores en dicho artículo para el caso de fidelidad cuadrática.

Introducción

En los últimos años, el campo de los problemas inversos ha sido sin duda una de las áreas de mayor crecimiento de la Matemática Aplicada. Este crecimiento se debe, en buena parte, a las aplicaciones en otras ciencias y en la industria.

En términos generales se dice que dos problemas son inversos el uno del otro si la formulación de uno involucra al otro ([30]). Generalmente por razones históricas suele llamarse a uno de ellos el problema *directo* (usualmente este es el más simple o el que se ha estudiado primero) y al otro, el problema *inverso*. Sin embargo, si detrás de un problema matemático existe un problema de la “vida real”, hay, en la gran mayoría de los casos, una distinción natural entre el problema directo y el problema inverso. Por ejemplo, si lo que se busca es predecir el comportamiento futuro de un sistema físico a partir del conocimiento de su estado presente y de las leyes físicas que lo gobiernan, entonces este es, naturalmente, el problema directo. En este caso, entre los posibles problemas inversos podemos mencionar la determinación del estado presente del sistema a partir de observaciones futuras (es decir la evolución “hacia atrás” en el tiempo) y la identificación de parámetros físicos a partir de observaciones de la evolución del sistema. Así, la resolución de un problema directo involucra análisis y razonamiento progresivos: de premisas a conclusiones, de causas a efectos, en ese orden; mientras que la resolución de un problema inverso, involucra un razonamiento regresivo: de conclusiones a premisas, de efectos a causas, en ese orden “inverso”.

En muchas ocasiones, un problema inverso resultar estar “mal condicionado” en el sentido de Hadamard, y pequeños errores en el dato pueden resultar en errores arbitrariamente grandes en la variable incógnita ([16]). Esta falta de dependencia continua de los datos se traduce en inestabilidad desde el punto de vista de la resolución numérica del problema, convirtiendo en inadecuados a los métodos numéricos estándar de aproximación desde el punto de vista práctico. Es precisamente esta pérdida de estabilidad la que origina la necesidad del desarrollo de herramientas y métodos matemáticos apropiados que permitan aproximar las soluciones del problema manteniendo la estabilidad del mismo, es decir, herramientas diseñadas para restaurar el proceso de inversión. En el contexto del estudio de problemas inversos, a tales herramientas y métodos se los denomina “métodos de regularización” ([16], [50]).

Uno de los métodos de regularización más conocidos, y sin dudas el más utilizado, es el método de Tikhonov-Phillips, que fue propuesto independientemente por D. L. Phillips y A. N. Tikhonov en 1962 y 1963, respectivamente ([43], [52], [53]). Si bien este método puede ser formulado en un contexto muy general mediante la utilización de teoría espectral ([16]), una de las razones por las cuales su aplicación es tan frecuente es, sin duda, el hecho de que también puede ser planteado de manera muy simple como un problema de optimización. Sin embargo, una de las principales restricciones del método de Tikhonov-Phillips, la que se encuentra ampliamente documentada en la literatura, es la relacionada a la *sobrerregularización* que produce su aplicación. Esta puede describirse de la siguiente manera: si la solución exacta posee discontinuidades o regiones en las cuales es localmente irregular (por ejemplo, existencia de vértices, saltos, bordes u otras características de falta de regularidad) entonces la aplicación de este método resulta en aproximaciones *sobresuavizadas*, las cuales pierden la mayor parte de la información en relación a tal falta de regularidad ([32]). Una manera de “subsananar” esto es a través de la utilización de diversos tipos de penalizantes que permiten capturar diferentes propiedades de la solución exacta.

En los últimos 25 años varios autores ([1], [24], [34], [35], [36], [48], [49], [50], [54]) han propuesto diversas generalizaciones del método clásico de regularización de Tikhonov-Phillips. La mayoría de ellas consisten, esencialmente, en considerar diferentes penalizantes. De esta manera, es posible pensar en funcionales de Tikhonov-Phillips con penalizante generalizado. En muchas aplicaciones resulta de gran utilidad además, el uso de otros términos de fidelidad diferentes del clásico cuadrático. Por ejemplo, la utilización de normas más débiles como término de fidelidad permite, en el caso de problemas de restauración de imágenes, una separación “más adecuada” de la componente de alta frecuencia, como lo son el ruido y la textura, de la parte suave de la imagen ([41], [42]). En ocasiones donde es sabido que el dato no es suave, la utilización de un término de fidelidad dado por $\|\cdot\|_1$ ha sido de gran utilidad ([11]). Por otro lado, en aplicaciones como en ejemplos de emisión de rayos X y tomografías computadas, es usual utilizar un término de fidelidad dado por el funcional de Kullback-Leibler ([44], [45]). De esta manera, es posible pensar en un método de Tikhonov-Phillips que sea generalizado tanto en el penalizante como en el término de fidelidad. Asimismo, un aspecto muy importante que debe considerarse en la aplicación de cualquier método de regularización es el de la elección del parámetro de regularización. En el contexto del método de Tikhonov-Phillips este parámetro establece un “compromiso” entre el término de fidelidad y el penalizante.

Esta tesis consta de tres capítulos. En el Capítulo 1, se aborda el estudio de la regularización de un problema inverso mal condicionado desde el enfoque determinístico de su resolución y en este contexto se presentan los métodos de Tikhonov-Phillips en sus versiones clásica, generalizada (es decir, con penalizante generalizado y fidelidad cuadrática) y doblemente generalizada (es decir, con fidelidad y penalizante generalizados). Mostraremos que estos métodos constituyen regularizaciones en el contexto de espacios de Banach. Para el caso de funcionales de tipo Tikhonov-Phillips generalizado, en la Sección 1.2, enunciamos los principales resultados sobre existencia y unicidad de soluciones regularizadas obtenidos por G. Mazzieri, R. Spies y K. Temperini en [34]. En la Sección 1.3, presentaremos generalizaciones de estos resultados que hemos obtenido para funcionales de tipo Tikhonov-Phillips doblemente generalizados. A su vez, para estos funcionales probaremos la existencia de solución de mínimo penalizante y la convergencia de las soluciones regularizadas a ella cuando el nivel de ruido tiende a cero por derecha y el parámetro de regularización es elegido mediante una regla a-priori.

Todo lo mencionado hasta aquí es de naturaleza absolutamente determinística, sin embargo, existe además un enfoque estadístico para resolver problemas inversos mal condicionados. La filosofía detrás de dicho enfoque está basada en la “reformulación” del problema inverso en una forma estadística de búsqueda de información; es decir, en un problema de inferencia. Así, los procedimientos de inversión estadística se basan en los siguientes principios consistentes con el enfoque Bayesiano:

- todas las variables incluidas en el modelo son consideradas aleatorias;
- el grado de información (o falta de ella) concerniente a esos valores se codifica en la distribución de probabilidades;
- la solución del problema inverso es una distribución de probabilidad llamada distribución a-posteriori.

Este último hecho hace que el enfoque estadístico para tratar, y resolver, un problema inverso sea muy diferente del enfoque clásico determinístico. Los métodos clásicos de regularización, producen una única estimación de la solución del problema, mientras que los métodos estadísticos dan como resultado una distribución de probabilidades que puede usarse para obtener, luego, estimaciones de la incógnita. Por lo tanto, una pregunta natural en este caso, no es cuál es el valor de la variable en cuestión (incógnita), sino *cuál es la información de la que se dispone acerca de esa variable*.

El enfoque estadístico Bayesiano es particularmente apropiado en problemas de restauración de imágenes porque permite la utilización de información cualitativa sobre la imagen que se

desea reconstruir o recuperar. Los primeros trabajos sobre la utilización de técnicas Bayesianas en el tratamiento de problemas inversos se deben a J. Besag ([3], [4]) y a S. y D. Geman ([19]), datan de mediados de la década del 80 y fueron precisamente aplicaciones al procesamiento y restauración de imágenes. La calidad de los resultados obtenidos desde entonces en problemas de restauración de imágenes ([6], [7]) y procesamiento de señales en general ([8]) son una clara demostración de la potencia de este nuevo enfoque.

Es importante mencionar que la definición abstracta de la solución “estadística” de un problema inverso, como la distribución a-posteriori de la incógnita, no es muy útil en la práctica si no se dispone de medios para explorarla. Los estimadores usualmente considerados son el estimador *máximo a-posteriori* y la *media condicional*.

Como se señaló anteriormente el enfoque estadístico Bayesiano parece estar, en principio, completamente desconectado de la teoría clásica que es puramente determinística. Sin embargo, los dos enfoques no son disjuntos y resulta de fundamental importancia la relación existente entre las distribuciones utilizadas en el enfoque Bayesiano y los funcionales presentes en el término de fidelidad y en el penalizante de la teoría generalizada de Tikhonov-Phillips.

En el Capítulo 2 de esta tesis estudiaremos la resolución de los problemas inversos desde el enfoque estadístico y probaremos que, bajo el supuesto de que la variable asociada al ruido y la incógnita tienen densidades pertenecientes a una familia exponencial, el estimador máximo a-posteriori coincide con la solución regularizada obtenida mediante la minimización de un funcional de tipo Tikhonov-Phillips generalizado o doblemente generalizado. Esto constituye el principal vínculo entre ambos enfoques de resolución de un problema inverso.

Otro aspecto muy importante que debe tenerse en consideración para obtener soluciones regularizadas es el de la adecuada elección del parámetro de regularización. Los métodos para la elección de este parámetro se clasifican en *a-priori*, si solo dependen del nivel de ruido, y *a-posteriori*, si dependen tanto del nivel de ruido como del dato. Existe, además, otro tipo de reglas para la elección de este parámetro que solo dependen del dato y se denominan *reglas heurísticas*. Entre ellas, podemos mencionar “el criterio de la curva L” propuesto en 1992 por P. C. Hansen ([22]) y el método de “validación cruzada generalizada” introducido en 1979 por G. H. Golub, M. T. Heat y G. Wahba ([20]). En 2011, K. Ito, B. Jin y J. Zou ([27]) propusieron una regla de elección del parámetro heurística para la regularización de un problema inverso mediante la minimización de un funcional de tipo Tikhonov-Phillips doblemente generalizado. A su vez, establecieron estimaciones del error a-posteriori cuando el término de fidelidad es de

tipo cuadrático y el penalizante es un funcional convexo ([25], [27], [26]). Sin embargo, para términos de fidelidad más generales no se conocen resultados de este tipo.

En el Capítulo 2 veremos cómo se construye esta regla heurística desde el enfoque estocástico de la resolución de un problema inverso a través de la formulación de modelos jerárquicos. Si bien no es un objetivo primordial de este trabajo de tesis el abordaje de una aplicación concreta, se considera que los resultados obtenidos son potencialmente transferibles y es por ello que implementaremos la regla estudiada en algunos ejemplos asociados al procesamiento de señales e imágenes.

Un resultado debido a A. B. Bakushinskii ([2]) muestra que un método de regularización no puede ser convergente cuando la regla de elección del parámetro asociada depende sólo del dato del problema. Por esta razón, dado que no se pueden obtener resultados de convergencia, en el Capítulo 3 presentaremos estimaciones del error a-posteriori que hemos obtenido para el caso en que la solución regularizada es obtenida mediante la minimización de funcionales de tipo Tikhonov-Phillips doblemente generalizados. Esta estimación se particulariza a la regla heurística dada en [27] dejando en evidencia que resulta mejor que la dada por los autores en dicho artículo para el caso de fidelidad cuadrática.

Los primeros resultados obtenidos en este capítulo, se basan principalmente en un supuesto dado por una condición fuente. Sin embargo, veremos que es posible reemplazar este supuesto por una desigualdad variacional obteniendo resultados análogos de estimaciones del error a-posteriori. Finalmente, se establecen relaciones entre la condición de optimalidad asociada al problema, la condición fuente y la desigualdad variacional mencionada.

Aunque la mayoría de los problemas que serán abordados y estudiados en esta tesis son de naturaleza predominantemente teórica, todos ellos proceden del estudio de problemas inversos mal condicionados que, como es sabido, aparecen en muchas áreas de la ciencia y, por lo tanto, es importante señalar que algunos resultados podrían tener potenciales aplicaciones en las mismas.

Con el objetivo de relacionar los contenidos abordados en cada capítulo, al final de cada uno de ellos se presentará un cuadro integrador. Y, luego de las conclusiones, un cuadro final que permite una visualización global de todos los contenidos y resultados obtenidos en este trabajo de tesis.

CAPÍTULO 1

Problemas inversos desde el enfoque determinístico

En este capítulo presentaremos algunas nociones básicas para la resolución de un problema inverso desde el enfoque determinístico clásico.

En un contexto general, un problema inverso puede formularse como la necesidad de determinar x en una ecuación de la forma

$$\mathcal{T}x = y, \tag{1.1}$$

donde $\mathcal{T} : \mathcal{X} \rightarrow \mathcal{Y}$ es un operador lineal y acotado entre dos espacios de Banach de dimensión infinita ¹ (en general, espacios de funciones) e y es el dato que se supone conocido, quizás con algún nivel de error.

Muy a menudo, los problemas inversos resultan *mal condicionados* (“ill-posed”) en el sentido de Hadamard, es decir, no verifican al menos uno de los siguientes postulados de Hadamard de *buen condicionamiento* (“well-posedness”):

(Had 1) Existencia: Para todo dato admisible, existe una solución.

(Had 2) Unicidad: Para todo dato admisible, la solución es única.

(Had 3) Dependencia continua: La solución depende del dato de manera continua.

Problemas de este tipo se presentan en diversas áreas y en una gran variedad de aplicaciones. Por ejemplo, la determinación de la distribución de la densidad en un cuerpo mediante la emisión de rayos X , que constituye la base de la Tomografía Computada ([29], [38]), se realiza mediante la inversión de la Transformada de Radon, resultando siempre un problema mal condicionado ([16]). Asimismo, en algunos problemas tales como la eliminación del ruido o la restauración de señales e imágenes ([31]), el modelo matemático da lugar a una integral de convolución con un núcleo de dispersión, generando ecuaciones integrales de primera clase, cuya inversión, salvo en el caso de núcleos degenerados, es siempre mal condicionada ([16]).

Es sabido que si el postulado (Had 1) no se verifica, es posible *relajar* la noción de solución, aceptando una solución generalizada del problema, al menos para dato exacto. Por otro lado, si el postulado (Had 2) es el que no se satisface, puede buscarse aquella solución que cumpla

¹En la teoría más clásica los espacios \mathcal{X} e \mathcal{Y} son espacios de Hilbert.

con alguna condición adicional, como por ejemplo, la de mínima norma en \mathcal{X} . Finalmente, la violación del tercer postulado (Had 3) es, en general, la principal causa de mal condicionamiento de un problema inverso. En este caso, pequeños errores o ruidos en la medición del dato y pueden producir errores muy grandes en la solución del problema inverso, volviendo inestables todos los procedimientos numéricos tradicionales de aproximación. Para restablecer la estabilidad es necesario implementar *métodos de regularización* cuya definición se introducirá en el contexto de espacios de Hilbert en la Sección 1.1, y en la Sección 1.2 se presentará una generalización de este concepto en espacios de Banach. Esta última definición será con la que se trabajará a lo largo de esta tesis.

Ahora veremos cómo se relacionan los postulados de Hadamard con el operador \mathcal{T} asociado al problema. Se dice que el dato $y \in \mathcal{Y}$ es *alcanzable* si $y \in \mathcal{R}(\mathcal{T})$. De esta manera, el postulado (Had 1) es equivalente a que cada $y \in \mathcal{Y}$ sea alcanzable, es decir que $\mathcal{R}(\mathcal{T}) = \mathcal{Y}$. Por otro lado, el postulado (Had 2) se verifica si y sólo si $\mathcal{N}(\mathcal{T}) = \{0\}$. Por lo tanto, que se verifiquen los postulados (Had 1) y (Had 2) es equivalente a que exista el operador \mathcal{T}^{-1} . Finalmente, que el postulado (Had 3) se cumpla es equivalente a que el operador \mathcal{T}^{-1} sea continuo (o acotado). Pero considerar que $\mathcal{R}(\mathcal{T}) = \mathcal{Y}$ y que $\mathcal{N}(\mathcal{T}) = \{0\}$ es muy restrictivo. Incluso si y no es alcanzable, quizás estamos interesados en encontrar alguna solución generalizada del problema (1.1), es decir, alguna solución que satisfaga (1.1) de manera aproximada. Además, si $\mathcal{N}(\mathcal{T}) \neq \{0\}$ la solución de (1.1) resulta no ser única y podemos elegir alguna que satisfaga una condición adicional, como señalamos anteriormente. Esta noción generalizada de solución será proporcionada, para el caso de espacios de Hilbert, a partir del concepto de inversa generalizada de Moore-Penrose del operador \mathcal{T} . En la Sección 1.1 presentaremos esta inversa y algunas de sus propiedades más importantes. La continuidad de la inversa generalizada de Moore-Penrose será relevante cuando el postulado (Had 3) no se cumpla.

Como hemos mencionado anteriormente, en la Sección 1.1 veremos en qué consisten los métodos de regularización cuando el operador asociado al problema está definido entre espacios de Hilbert, haciendo especial hincapié en el método de regularización de Tikhonov-Phillips. Una de las ventajas de este último método es su formulación como un problema de optimización en el cual la solución regularizada viene dada por un minimizante del funcional de Tikhonov-Phillips clásico (que llamaremos “funcional TP”) dado por

$$\mathcal{J}_\eta(x) = \|\mathcal{T}x - y\|^2 + \eta \|x\|^2,$$

con parámetro $\eta > 0$. Luego, en la Sección 1.2 generalizaremos el concepto de regularización en el contexto de espacios de Banach y estudiaremos el método que consiste en aproximar la solución del problema (1.1) mediante un minimizante del funcional de Tikhonov-Phillips con penalizante generalizado

$$\mathcal{J}_{\psi,\eta}(x) = \|\mathcal{T}x - y\|^2 + \eta\psi(x),$$

donde ψ es un funcional real definido en el espacio \mathcal{X} . A este funcional lo llamaremos “funcional TPG” y, para este método, enunciaremos los principales resultados de existencia y unicidad de minimizante que pueden encontrarse en [34]. Finalmente, en la Sección 1.3 presentaremos generalizaciones de estos resultados que hemos obtenido para el caso del funcional de Tikhonov-Phillips con término de fidelidad y penalizante generalizados dado por

$$\mathcal{J}_{\phi,\psi,\eta}(x) = \phi(\mathcal{T}x, y) + \eta\psi(x),$$

donde ϕ es un funcional real definido sobre el espacio $\mathcal{Y} \times \mathcal{Y}$. Como el funcional $\mathcal{J}_{\phi,\psi,\eta}$ es generalizado en sus dos términos se denotará como “funcional TPGG”. A su vez, enunciaremos y probaremos resultados de convergencia de soluciones regularizadas obtenidas mediante la minimización del funcional TPGG para reglas de elección del parámetro a-priori.

1.1. Teoría Clásica

A lo largo de esta sección presentaremos el concepto de inversa generalizada de Moore-Penrose del operador \mathcal{T} definido entre dos espacios de Hilbert, enunciaremos algunas propiedades importantes de la misma así como condiciones necesarias y suficientes para su continuidad. Para mayor detalle sobre estos resultados, el lector interesado puede consultar [16]. Finalmente, veremos en qué consisten los métodos de regularización y, en particular, nos abocaremos al estudio del método de Tikhonov-Phillips en su versión clásica.

En primer lugar, presentamos los conceptos de solución de mínimos cuadrados y de mejor solución aproximada.

Definición 1.1. Sean \mathcal{X} e \mathcal{Y} dos espacios de Hilbert y $\mathcal{T} : \mathcal{X} \rightarrow \mathcal{Y}$ un operador lineal y acotado.

1. $x_0 \in \mathcal{X}$ se dice *solución de mínimos cuadrados* de $\mathcal{T}x = y$ si

$$\|\mathcal{T}x_0 - y\| = \inf \{ \|\mathcal{T}z - y\| : z \in \mathcal{X} \}.$$

2. $x_0 \in \mathcal{X}$ se denomina *mejor solución aproximada* de $\mathcal{T}x = y$ si x_0 es una solución de mínimos cuadrados de $\mathcal{T}x = y$ y

$$\|x_0\| = \inf \{\|z\| : z \text{ es una solución de mínimos cuadrados de } \mathcal{T}x = y\}.$$

De la definición anterior se sigue inmediatamente que la mejor solución aproximada de (1.1) es la solución de mínimos cuadrados de mínima norma en \mathcal{X} .

A continuación, definimos la inversa generalizada de Moore-Penrose del operador \mathcal{T} . Para ello, se restringe su dominio y rango de tal manera que el operador restringido resultante $\tilde{\mathcal{T}}$ sea invertible.

Definición 1.2. Sean \mathcal{X} e \mathcal{Y} dos espacios de Hilbert, $\mathcal{T} \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ y $\tilde{\mathcal{T}} \doteq \mathcal{T}|_{\mathcal{N}(\mathcal{T})^\perp} : \mathcal{N}(\mathcal{T})^\perp \rightarrow \mathcal{R}(\mathcal{T})$. La *inversa generalizada de Moore-Penrose* \mathcal{T}^\dagger de \mathcal{T} se define como la única extensión lineal de $\tilde{\mathcal{T}}^{-1}$ al dominio

$$\mathcal{D}(\mathcal{T}^\dagger) \doteq \mathcal{R}(\mathcal{T}) \oplus \mathcal{R}(\mathcal{T})^\perp,$$

que satisface $\mathcal{N}(\mathcal{T}^\dagger) = \mathcal{R}(\mathcal{T})^\perp$.

En esta definición, se extiende $\tilde{\mathcal{T}}^{-1} : \mathcal{R}(\mathcal{T}) \rightarrow \mathcal{N}(\mathcal{T})^\perp$ a $\mathcal{T}^\dagger : \mathcal{R}(\mathcal{T}) \oplus \mathcal{R}(\mathcal{T})^\perp \rightarrow \mathcal{N}(\mathcal{T})^\perp$ de tal manera que $\mathcal{T}^\dagger y = \tilde{\mathcal{T}}^{-1} y_1$, donde $y = y_1 + y_2$ con $y_1 \in \mathcal{R}(\mathcal{T})$ y $y_2 \in \mathcal{R}(\mathcal{T})^\perp$. Por otra parte, es importante observar que, como $\mathcal{N}(\tilde{\mathcal{T}}) = \{0\}$ y $\mathcal{R}(\tilde{\mathcal{T}}) = \mathcal{R}(\mathcal{T})$ se sigue inmediatamente que $\tilde{\mathcal{T}}^{-1}$ existe y, como consecuencia, el operador \mathcal{T}^\dagger está bien definido.

Enunciamos ahora algunas propiedades importantes del operador \mathcal{T}^\dagger . La siguiente proposición presenta cuatro ecuaciones que caracterizan unívocamente a \mathcal{T}^\dagger .

Proposición 1.3. Sean P y Q las proyecciones ortogonales sobre $\mathcal{N}(\mathcal{T})$ y $\overline{\mathcal{R}(\mathcal{T})}$, respectivamente. Entonces, $\mathcal{R}(\mathcal{T}^\dagger) = \mathcal{N}(\mathcal{T})^\perp$ y se satisfacen las cuatro ecuaciones de Moore-Penrose:

$$\mathcal{T}\mathcal{T}^\dagger\mathcal{T} = \mathcal{T}$$

$$\mathcal{T}^\dagger\mathcal{T}\mathcal{T}^\dagger = \mathcal{T}^\dagger$$

$$\mathcal{T}^\dagger\mathcal{T} = I - P$$

$$\mathcal{T}\mathcal{T}^\dagger = Q|_{\mathcal{D}(\mathcal{T}^\dagger)}.$$

A continuación, presentamos una condición necesaria y suficiente para que un problema inverso sea bien condicionado en el sentido de Hadamard.

Proposición 1.4. *La inversa generalizada de Moore-Penrose \mathcal{T}^\dagger de \mathcal{T} es acotada si y sólo si $\mathcal{R}(\mathcal{T})$ es cerrado.*

El siguiente teorema afirma que la inversa generalizada de Moore-Penrose de \mathcal{T} es el operador solución que mapea el dato y sobre la mejor solución aproximada del problema $\mathcal{T}x = y$ (siempre que $y \in \mathcal{D}(\mathcal{T}^\dagger)$).

Teorema 1.5. *Si $y \in \mathcal{D}(\mathcal{T}^\dagger)$ entonces $\mathcal{T}x = y$ tiene una única mejor solución aproximada dada por $x^\dagger \doteq \mathcal{T}^\dagger y$. Además, el conjunto de todas las soluciones de mínimos cuadrados está dado por $S = \{x^\dagger + \tilde{x}, \tilde{x} \in \mathcal{N}(\mathcal{T})\}$.*

Como hemos mencionado previamente, cuando un problema inverso no verifica el postulado (Had 3), es necesario implementar un método de regularización. En términos generales, la regularización de un problema inverso mal condicionado consiste en su aproximación mediante una familia, convenientemente construída, de problemas inversos “vecinos” y bien condicionados. Sin embargo, se debe tener presente que ningún truco matemático puede hacer estable un problema que es intrínsecamente inestable. Todo lo que un método de regularización puede hacer es recuperar información parcial acerca de la solución, de modo tan estable como sea posible. De esta manera, dado un dato $y \in \mathcal{D}(\mathcal{T}^\dagger)$, se desea hallar la mejor solución aproximada $x^\dagger = \mathcal{T}^\dagger y$ de (1.1) a la que llamaremos *solución exacta*. En general, el dato exacto y no es conocido y sólo se dispone de una aproximación o *dato con ruido*, que denotaremos y^δ , el cual satisface

$$\|y^\delta - y\| \leq \delta, \quad (1.2)$$

donde $\delta > 0$ es el *nivel de ruido* asociado a los errores de medición.

Es claro que, para problemas inversos mal condicionados, $\mathcal{T}^\dagger y^\delta$ (si existe, pues en este caso $\mathcal{D}(\mathcal{T}^\dagger)$ es un subespacio propio de \mathcal{Y}) no constituye una buena aproximación de $\mathcal{T}^\dagger y$ debido a la falta de continuidad del operador \mathcal{T}^\dagger . Como consecuencia, se desea encontrar una aproximación de x^\dagger , denotada por x_η^δ , que dependa de manera continua del dato con ruido y^δ , que pueda calcularse de manera estable y que, finalmente, tienda a x^\dagger a medida que el nivel de ruido δ tiende a cero por derecha, siempre que el *parámetro de regularización* η sea elegido de manera “apropiada”.

En este sentido, el problema (1.1) no es pensado sólo para un dato específico y sino que se considera como una colección de problemas para cada $y \in \mathcal{R}(\mathcal{T})$ (si se supone alcanzabilidad) o para cada $y \in \mathcal{D}(\mathcal{T}^\dagger)$. De esta manera, no se regulariza un problema en particular sino una colección de problemas o, en otras palabras, regularizamos el operador \mathcal{T}^\dagger . Más precisamente,

una regularización de \mathcal{T}^\dagger consiste en una familia de operadores continuos $\{R_\eta\}$ (dependiente de un parámetro η) mediante la cual se aproxima la solución x^\dagger por $x_\eta^\delta \doteq R_\eta y^\delta$, calculada en forma estable (debido a la continuidad de R_η). Por otro lado, es sabido que el parámetro η , vinculado a δ y/o a y^δ y, quizás, en menor medida a \mathcal{T} o a y , debe ser elegido de manera tal que x_η^δ tienda a x^\dagger a medida que el nivel de ruido δ tienda a cero por derecha. Definimos ahora un método de regularización.

Definición 1.6. Sean $\mathcal{T} : \mathcal{X} \rightarrow \mathcal{Y}$ un operador lineal y acotado entre dos espacios de Hilbert y $\eta_0 \in (0, +\infty]$. Para cada $\eta \in (0, \eta_0)$, sea $R_\eta : \mathcal{Y} \rightarrow \mathcal{X}$ un operador continuo (no necesariamente lineal). La familia $\{R_\eta\}$ se denomina *familia de operadores de regularización para \mathcal{T}^\dagger* si, para cada $y \in \mathcal{D}(\mathcal{T}^\dagger)$, existe una *regla de elección del parámetro* $\eta = \eta(\delta, y^\delta)$ tal que

$$\lim_{\delta \rightarrow 0^+} \sup_{y^\delta \in \mathcal{Y}: \|y - y^\delta\| \leq \delta} \|R_{\eta(\delta, y^\delta)} y^\delta - \mathcal{T}^\dagger y\| = 0. \quad (1.3)$$

Además, $\eta : \mathbb{R}^+ \times \mathcal{Y} \rightarrow (0, \eta_0)$ es tal que

$$\lim_{\delta \rightarrow 0^+} \sup_{y^\delta \in \mathcal{Y}: \|y - y^\delta\| \leq \delta} \eta(\delta, y^\delta) = 0. \quad (1.4)$$

Para $y \in \mathcal{D}(\mathcal{T}^\dagger)$ fijo, el par $(\{R_\eta\}, \eta(\delta, y^\delta))$ se denomina *método de regularización convergente* (para resolver (1.1)) si se satisfacen (1.3) y (1.4).

Así, un método de regularización consiste en una familia de operadores de regularización (que podrían ser no lineales) y una regla de elección del parámetro. Es oportuno observar que una familia de operadores $\{R_\eta\}$ se define para \mathcal{T}^\dagger mientras que una regla de elección del parámetro η se define para un dato específico y . En la siguiente definición se presentan distintos tipos de reglas de elección del parámetro.

Definición 1.7. Sea η una regla de elección del parámetro. Si η no depende de y^δ pero sí de δ , se denomina *regla de elección del parámetro a-priori* (debido a que puede definirse antes de la obtención del dato y^δ) y se denota $\eta = \eta(\delta)$. De lo contrario, se denomina *regla de elección del parámetro a-posteriori* y se denota con $\eta = \eta(\delta, y^\delta)$.

El siguiente teorema debido a A. B. Bakushinskii ([2]) muestra que una regla de elección del parámetro que sólo dependa de y^δ (es decir, que no dependa del nivel de ruido δ) no puede ser parte de un método de regularización convergente para un problema inverso mal condicionado en el sentido de la Definición 1.6. Este tipo de regla se denomina *regla de elección del parámetro heurística*. Entre ellas, podemos mencionar “el criterio de la curva L” propuesto en 1992 por P.

C. Hansen ([22]) y el método de “validación cruzada generalizada” introducido en 1979 por G. H. Golub, M. T. Heat y G. Wahba ([20]). En el Capítulo 2 de esta tesis veremos en detalle una regla de este tipo propuesta en 2011 por K. Ito, B. Jin y J. Zou ([27]).

Teorema 1.8. *Sea $\mathcal{T} : \mathcal{X} \rightarrow \mathcal{Y}$ un operador lineal y acotado entre dos espacios de Hilbert. Supongamos además que existen una familia de operadores de regularización $\{R_\eta\}$ para \mathcal{T}^\dagger y una regla de elección del parámetro η que depende sólo de y^δ pero no de δ , tales que el método de regularización $(\{R_\eta\}, \eta(y^\delta))$ es convergente para cada $y \in \mathcal{D}(\mathcal{T}^\dagger)$. Entonces, \mathcal{T}^\dagger es acotado (y luego, el problema (1.1) es bien condicionado).*

Es importante observar que este resultado no dice que si \mathcal{T}^\dagger no es acotado, una regla de elección del parámetro que no depende de δ (conocida como *libre de error*) no pueda tener un buen comportamiento para niveles de ruido δ pequeños. Además, cabe mencionar aquí que el Teorema 1.8 es un corolario del resultado presentado por A. B. Bakushinskii en su artículo original ([2]) donde el operador \mathcal{T} está definido entre dos espacios métricos.

La proposición a continuación permite afirmar que se pueden construir operadores de regularización como aproximaciones puntuales de la inversa generalizada de Moore-Penrose de \mathcal{T} .

Proposición 1.9. *Sean $\mathcal{T} : \mathcal{X} \rightarrow \mathcal{Y}$ un operador lineal y acotado entre dos espacios de Hilbert y R_η un operador continuo (no necesariamente lineal) para cada $\eta > 0$. Entonces, la familia $\{R_\eta\}$ es una familia de operadores de regularización para \mathcal{T}^\dagger si $R_\eta y \rightarrow \mathcal{T}^\dagger y$ para todo $y \in \mathcal{D}(\mathcal{T}^\dagger)$, cuando $\eta \rightarrow 0^+$. En este caso, para cada $y \in \mathcal{D}(\mathcal{T}^\dagger)$ existe una regla de elección del parámetro a-priori, $\eta(\delta)$, tal que $(\{R_\eta\}, \eta(\delta))$ es un método de regularización convergente para resolver $\mathcal{T}x = y$.*

En el siguiente resultado se caracterizan las reglas de elección del parámetro a-priori que originan métodos de regularización convergentes.

Proposición 1.10. *Sean $\mathcal{T} : \mathcal{X} \rightarrow \mathcal{Y}$ un operador lineal y acotado entre dos espacios de Hilbert y $\{R_\eta\}$ una familia de operadores lineales de regularización; para cada $y \in \mathcal{D}(\mathcal{T}^\dagger)$ sea $\eta : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ una regla de elección del parámetro a-priori. Entonces, $(\{R_\eta\}, \eta(\delta))$ es un método de regularización convergente si y sólo si*

$$\lim_{\delta \rightarrow 0^+} \eta(\delta) = 0 \tag{1.5}$$

y

$$\lim_{\delta \rightarrow 0^+} \delta \|R_{\eta(\delta)}\| = 0. \tag{1.6}$$

Si bien existe una gran variedad de métodos de regularización tales como el método de Landweber, la descomposición en valores singulares truncada, el método de Showalter ([16]), uno de los más conocidos y, sin dudas, el más utilizado es el método de regularización de Tikhonov-Phillips, propuesto independientemente por D. L. Phillips y A. N. Tikhonov en 1962 y 1963, respectivamente (ver [43], [52], [53]). El mismo puede ser formulado en un contexto muy general mediante la utilización de teoría espectral ([16]) en el cual la familia de operadores de regularización $\{R_\eta\}$ está dada por $R_\eta = (\mathcal{T}^* \mathcal{T} + \eta I)^{-1} \mathcal{T}^*$. Luego, si sólo se dispone del dato con ruido y^δ , la solución regularizada del problema (1.1) obtenida por este método resulta

$$x_\eta^\delta = (\mathcal{T}^* \mathcal{T} + \eta I)^{-1} \mathcal{T}^* y^\delta, \quad (1.7)$$

donde $\eta > 0$ es un parámetro de regularización fijo.

Una de las razones por las cuales la aplicación de este método es tan frecuente es, sin dudas, el hecho de que también puede ser planteado de manera muy simple como un problema de optimización. En efecto, para cada parámetro de regularización $\eta > 0$, la solución regularizada x_η^δ dada en (1.7) coincide con el único minimizante del funcional TP dado por

$$\mathcal{J}_\eta(x) = \|\mathcal{T}x - y^\delta\|^2 + \eta \|x\|^2, \quad (1.8)$$

como se demuestra en el Teorema 5.1 de [16]. El término $\|\mathcal{T}x - y^\delta\|^2$ se conoce como *término de fidelidad* y mide la proximidad entre $\mathcal{T}x$ y y^δ . Por otro lado, $\|x\|^2$ se denomina *funcional de penalización* o *penalizante*. La minimización de (1.8) es un “compromiso” entre minimizar la norma del residuo, $\mathcal{T}x - y^\delta$, y mantener el término de penalización $\|x\|^2$ pequeño, es decir, forzar estabilidad. Es importante señalar aquí que el penalizante en (1.8) no sólo introduce estabilidad sino que también determina ciertas propiedades de regularidad de las soluciones aproximadas x_η^δ , y de la correspondiente solución de mínimos cuadrados a la que tienden, cuando el parámetro de regularización tiende a cero por derecha. El parámetro de regularización η determina un “trade-off” entre el término de fidelidad y el penalizante. Valores grandes del parámetro η corresponden a una mayor estabilidad de las soluciones aproximadas, pero el problema auxiliar no resulta ser “vecino” ya que se encuentra alejado del original. Por otro lado, valores de η pequeños representan problemas auxiliares cercanos al original, pero tienden a volverse más inestables a medida que η se acerca a cero.

A continuación, enunciaremos un resultado de convergencia de las soluciones regularizadas obtenidas mediante el método de Tikhonov-Phillips cuando la regla de elección del parámetro es a-priori que resulta ser una versión particular de la Proposición 1.10 para este método.

Proposición 1.11. Sea x_η^δ definido como en (1.7), $y \in \mathcal{R}(\mathcal{T})$, $\|y^\delta - y\| \leq \delta$. Si la regla de elección del parámetro $\eta = \eta(\delta)$ es tal que

$$\lim_{\delta \rightarrow 0^+} \eta(\delta) = 0 \quad (1.9)$$

y

$$\lim_{\delta \rightarrow 0^+} \frac{\delta^2}{\eta(\delta)} = 0 \quad (1.10)$$

entonces $\lim_{\delta \rightarrow 0^+} x_{\eta(\delta)}^\delta = \mathcal{T}^\dagger y$.

La prueba de este resultado puede encontrarse en [16]. En la misma se utilizan argumentos de la teoría de optimización, es decir, el hecho de que x_η^δ es un minimizante del funcional \mathcal{J}_η . Sin embargo, este resultado puede probarse utilizando teoría espectral para mostrar que la hipótesis (1.10) de la Proposición 1.11 implica la hipótesis (1.6) de la Proposición 1.10 (notar que las hipótesis (1.5) y (1.9) coinciden).

Una de las principales restricciones del método de Tikhonov-Phillips, la que se encuentra ampliamente documentada en la literatura, es la relacionada a la *sobrerregularización* que produce su aplicación. Esta puede describirse de la siguiente manera: si la solución exacta posee discontinuidades o regiones en las cuales es localmente irregular (por ejemplo, existencia de vértices, saltos, bordes u otras características de falta de regularidad) entonces la aplicación de este método resulta en aproximaciones *sobresuavizadas*, las cuales pierden la mayor parte de la información en relación a tal falta de regularidad ([32]). En algunas aplicaciones concretas (tal es el caso, por ejemplo, en ciertos problemas asociados a restauración de imágenes en Astronomía y Medicina) en las cuales es muy importante preservar las características de tales irregularidades, esta particularidad del método de Tikhonov-Phillips puede ser altamente indeseable. Una manera de “subsana” esto es a través de la utilización de diversos tipos de penalizantes que permiten capturar diferentes propiedades de la solución exacta. La importancia de la utilización de este nuevo enfoque radica en el hecho que no solamente puede inducirse estabilidad con diferentes penalizantes sino que, además, una adecuada elección del mismo, basada quizás en el conocimiento de cierto tipo de información a-priori acerca de la solución exacta, puede resultar en soluciones aproximadas que compartan con aquella propiedades similares. De esta manera, se dio origen a una amplia variedad de métodos de Tikhonov-Phillips generalizados.

En sus artículos originales ([52], [53]) A. N. Tikhonov consideró un método de regularización con un funcional más general dado por

$$\mathcal{J}_{\mathcal{L},\eta}(x) = \|\mathcal{T}x - y\|^2 + \eta \|\mathcal{L}x\|^2, \quad (1.11)$$

donde $\mathcal{L} : \mathcal{D}(\mathcal{L}) \subset \mathcal{X} \rightarrow \mathcal{Z}$ es un operador a valores en un espacio de Hilbert \mathcal{Z} . En 1996, Engl et al. ([16]) probaron, bajo ciertas condiciones sobre el operador \mathcal{L} , la existencia de minimizantes del funcional (1.11) y la convergencia de éstos a la solución de mínimos cuadrados de (1.1) que además minimiza $\|x\|_{\mathcal{L}} \doteq \|\mathcal{L}x\|$. Sin embargo, si \mathcal{L} corresponde, por ejemplo, a un operador diferencial, las soluciones regularizadas obtenidas mediante la minimización del funcional (1.11) poseen la misma particularidad indeseable que aquellas obtenidas mediante la minimización del funcional TP.

En los últimos 25 años varios autores ([1], [24], [35], [36], [47], [48], [49], [50], [54]) han propuesto diversas generalizaciones del método clásico de regularización de Tikhonov-Phillips. La mayoría de ellas consisten, esencialmente, en considerar diferentes penalizantes. Según la elección del mismo, el método origina soluciones regularizadas con distintas propiedades, las que aproximan diferentes soluciones de mínimos cuadrados de (1.1). Por ejemplo, la utilización de la norma de variación acotada, $\|x\|_{\text{BV}}$, como penalizante en (1.11) en lugar de $\|\mathcal{L}x\|^2$, origina el llamado *método de regularización por variación acotada* ([1], [12], [47], [48]). Este método solamente impone que las soluciones sean de variación acotada y no necesariamente suaves. Por esta razón resulta muy apropiado en ciertos problemas de restauración de imágenes en los cuales es imprescindible que la regularización sea capaz de detectar y restaurar bordes y discontinuidades que se presumen presentes en la solución exacta ([10], [14]). Por ejemplo, en el área de Medicina son de particular interés los problemas relacionados con la detección precisa de regiones tumorales ([40]). De esta manera, es posible pensar en funcionales de Tikhonov-Phillips con penalizante generalizado que serán estudiados en detalle en la siguiente sección.

1.2. Regularización de tipo Tikhonov-Phillips generalizado

Como hemos mencionado, a lo largo de esta sección consideraremos funcionales TPG de la forma

$$\mathcal{J}_{\psi,\eta}(x) \doteq \|\mathcal{T}x - y\|^2 + \eta\psi(x),$$

donde $\mathcal{T} : \mathcal{X} \rightarrow \mathcal{Y}$ es un operador lineal y acotado entre dos espacios de Banach, el penalizante ψ es un funcional real arbitrario con dominio $D \subset \mathcal{X}$, $x \in D$ y η es una constante positiva.

Para el caso clásico estudiado en la sección anterior, donde \mathcal{X} e \mathcal{Y} eran espacios de Hilbert y el funcional ψ estaba dado por $\psi(x) = \|x\|^2$, hemos visto que las soluciones regularizadas convergen a la mejor solución aproximada, es decir, a aquella solución de mínima norma o de

mínimo penalizante. En este sentido, a continuación presentamos un concepto que extiende el de solución de mínima norma y que será de nuestro interés de aquí en adelante.

Definición 1.12. Decimos que $\tilde{x} \in \mathcal{X}$ es una *solución ψ -minimizante* del problema (1.1) si $\mathcal{T}\tilde{x} = y$ y si $\psi(\tilde{x}) = \inf \{\psi(\hat{x}) : \hat{x} \in \mathcal{X} \text{ y } \mathcal{T}\hat{x} = y\}$.

Como hemos mencionado, un problema inverso es mal condicionado si no verifica al menos uno de los tres postulados de Hadamard de buen condicionamiento. Hemos visto que si (Had 1) o (Had 2) no se verifican, es posible relajar el concepto de solución o considerar aquella solución que cumpla con alguna condición adicional, respectivamente. Además, para el caso en que \mathcal{X} e \mathcal{Y} son espacios de Hilbert, el cumplimiento del postulado (Had 3) se encuentra estrechamente relacionado con el hecho que el operador \mathcal{T}^\dagger sea acotado y, como consecuencia de la Proposición 1.4, con que el rango de \mathcal{T} sea cerrado. Cuando \mathcal{X} e \mathcal{Y} son espacios de Banach, el cumplimiento del postulado (Had 3) también está vinculado con el hecho que el rango del operador \mathcal{T} sea cerrado. Esto último refiere al concepto de buen condicionamiento según M. Z. Nashed ([37]), que presentamos a continuación.

Definición 1.13. El problema inverso (1.1) se dice *bien condicionado en el sentido de Nashed* si el rango del operador \mathcal{T} es cerrado, es decir, si $\mathcal{R}(\mathcal{T}) = \overline{\mathcal{R}(\mathcal{T})}$; y *mal condicionado en el sentido de Nashed* si el rango de \mathcal{T} no es cerrado.

Es posible probar que si el operador \mathcal{T} es inyectivo, el problema (1.1) es mal condicionado en el sentido de Nashed si y sólo si el operador $\mathcal{T}^{-1} : \mathcal{R}(\mathcal{T}) \subset \mathcal{Y} \rightarrow \mathcal{X}$ es no acotado, es decir, si el tercer postulado de Hadamard no se verifica ([50], Proposición 3.9). Así, se puede observar que, si el operador \mathcal{T} es inyectivo, los conceptos de mal condicionamiento de Hadamard y Nashed son equivalentes.

Como es sabido, cuando un problema inverso no verifica el postulado (Had 3), es necesario implementar un método de regularización. En la Sección 1.1, hemos estudiado en qué consisten estos métodos en el contexto de espacios de Hilbert. Sin embargo, es necesario generalizar este concepto cuando los espacios son de Banach.

Definición 1.14. Sean $\mathcal{T} : \mathcal{X} \rightarrow \mathcal{Y}$ un operador lineal y acotado entre dos espacios de Banach, $y \in \mathcal{R}(\mathcal{T})$ y $\bar{\eta} \in (0, +\infty)$. Un mapeo que transforma cada par $(y^\delta, \eta) \in \mathcal{Y} \times (0, \bar{\eta}]$ a algún elemento bien definido $x_\eta^\delta \in \mathcal{X}$ se denomina *regularización* del problema inverso (1.1) si existe una elección apropiada $\eta = \eta(y^\delta, \delta)$ del *parámetro de regularización* tal que, para cada sucesión $\{y_n\}_n \subset \mathcal{Y}$ con $\|y_n - y\| \leq \delta_n$, donde $\delta_n \rightarrow 0$ cuando $n \rightarrow \infty$, la *solución regularizada* correspondiente

$x_{\eta(y_n, \delta_n)}^{\delta_n}$ converge, en algún sentido, a la solución de (1.1). Si esta solución no es única, las soluciones regularizadas deben converger a una solución ψ -minimizante.

En la Sección 1.1 hemos presentado el método de Tikhonov-Phillips en su versión clásica el cual, como vimos, consiste en aproximar la solución del problema (1.1), en el contexto de espacios de Hilbert, mediante un minimizante del funcional TP dado en (1.8). Hemos visto que este método es un método de regularización convergente en el sentido de la Definición 1.6 bajo ciertas hipótesis sobre la regla de elección del parámetro. El objetivo principal de esta sección y de la siguiente es estudiar versiones generalizadas del método de Tikhonov-Phillips y determinar si constituyen regularizaciones en el sentido de la Definición 1.14. Proponemos entonces una primera generalización de este método, la cual consiste en aproximar la solución del problema (1.1) mediante un minimizante del funcional TPG definido como

$$\mathcal{J}_{\psi, \eta}(x) \doteq \|\mathcal{T}x - y\|^2 + \eta \psi(x), \quad (1.12)$$

donde $\mathcal{T} : \mathcal{X} \rightarrow \mathcal{Y}$ es un operador lineal y acotado entre dos espacios de Banach, el penalizante ψ es un funcional real arbitrario con dominio $D \subset \mathcal{X}$, $x \in D$ y η es una constante positiva. Más adelante, en la Sección 1.3, propondremos un método aún más general que consiste en aproximar la solución del problema (1.1) mediante un minimizante del funcional TPGG

$$\mathcal{J}_{\phi, \psi, \eta}(x) \doteq \phi(\mathcal{T}x, y) + \eta \psi(x),$$

donde ϕ es un funcional real definido sobre el espacio $\mathcal{Y} \times \mathcal{Y}$.

Para poder determinar si estos nuevos métodos propuestos constituyen una regularización en el sentido de la Definición 1.14 es necesario, en primer lugar, probar la existencia y unicidad de soluciones regularizadas, es decir, la existencia y unicidad de minimizante para los funcionales TPG y TPGG, para cada dato $y \in \mathcal{Y}$. En esta sección lo haremos para el caso del funcional TPG y, en la Sección 1.3, para el caso del funcional TPGG. Finalmente, será necesario probar la existencia de una solución ψ -minimizante y estudiar la convergencia, en algún sentido, de las soluciones regularizadas obtenidas a partir de estos métodos a dicha solución cuando el nivel de ruido tiende a cero. Esto lo haremos en la Sección 1.3 para el caso más general, es decir, para el funcional TPGG.

En [34] se determinaron condiciones suficientes sobre el penalizante ψ que garantizan existencia, unicidad y estabilidad fuerte y débil de minimizantes del funcional TPG bajo distintos tipos de perturbaciones. Antes de enunciar algunos de estos resultados, se introducen ciertas definiciones necesarias.

Definición 1.15. Sean \mathcal{X} un espacio vectorial, ψ un funcional definido sobre el conjunto $D \subset \mathcal{X}$ y A un subconjunto de D . Diremos que A es ψ -acotado si existe una constante $k < \infty$ tal que $|\psi(a)| \leq k$, para todo $a \in A$.

Definición 1.16. Sean \mathcal{X} un espacio vectorial y ψ, F dos funcionales definidos sobre un conjunto $D \subset \mathcal{X}$. Diremos que el funcional F es ψ -coercivo si $\lim_{n \rightarrow \infty} F(x_n) = +\infty$ para cada sucesión $\{x_n\} \subset D$ tal que $\lim_{n \rightarrow \infty} \psi(x_n) = +\infty$.

Observación 1.17. Si F es un funcional ψ -coercivo y ψ está acotado por abajo, entonces todos los conjuntos de nivel para F definidos por $S_a \doteq \{x \in D : F(x) \leq a\}$, con $a \in \mathbb{R}$ son conjuntos ψ -acotados. En efecto, sea $a \in \mathbb{R}$ y supongamos que el conjunto S_a no es ψ -acotado, entonces para cada $k \in \mathbb{N}$, existe $x_k \in S_a$ tal que $|\psi(x_k)| > k$. Por lo tanto, $\lim_{k \rightarrow \infty} \psi(x_k) = +\infty$. Luego, como F es ψ -coercivo, tenemos que $\lim_{k \rightarrow \infty} F(x_k) = +\infty$. Pero para cada $k \in \mathbb{N}$, $x_k \in S_a$ y, como consecuencia, $F(x_k) \leq a$ de donde resulta que no puede ocurrir que $\lim_{k \rightarrow \infty} F(x_k) = +\infty$. Puesto que esta contradicción provino de suponer que S_a no es ψ -acotado, se sigue que S_a es ψ -acotado para todo $a \in \mathbb{R}$.

Definición 1.18. Sean \mathcal{X} un espacio vectorial normado y F, ψ dos funcionales con $\mathcal{D}(F) \subset \mathcal{D}(\psi) \subset \mathcal{X}$. Diremos que F es ψ -subsecuencialmente (débilmente) semicontinuo inferiormente si para cada sucesión ψ -acotada $\{x_n\} \subset \mathcal{D}(F)$ tal que $x_n \xrightarrow{(\omega)} x \in \mathcal{D}(F)$, existe una subsucesión $\{x_{n_j}\} \subset \{x_n\}$ tal que $F(x) \leq \liminf_{j \rightarrow \infty} F(x_{n_j})$. Si F es ψ -subsecuencialmente semicontinuo inferiormente diremos simplemente que F es ψ -ssi. Análogamente, si F es ψ -subsecuencialmente débilmente semicontinuo inferiormente diremos que F es ψ -sdsi.

Los principales resultados sobre existencia y unicidad de un minimizante global para el funcional (1.12) se detallan a continuación.

Teorema 1.19. Sean \mathcal{X} un espacio vectorial normado, \mathcal{Y} un espacio con producto interno, $\mathcal{T} \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$, $y \in \mathcal{Y}$, $D \subset \mathcal{X}$ un conjunto convexo y $\psi : D \rightarrow \mathbb{R}$ un funcional que satisface las siguientes hipótesis:

(H1): está acotado por abajo, es decir, existe $\gamma \geq 0$ tal que $\psi(x) \geq -\gamma$, para todo $x \in D$;

(H2): para cada sucesión ψ -acotada $\{x_n\} \subset D$ tal que $x_n \xrightarrow{\omega} x \in D$, existe una subsucesión $\{x_{n_j}\} \subset \{x_n\}$ que verifica $\psi(x) \leq \liminf_{j \rightarrow \infty} \psi(x_{n_j})$;

(H3): las sucesiones ψ -acotadas son relativamente débilmente compactas en D , es decir, para cada sucesión ψ -acotada $\{x_n\} \subset D$, existen $x \in D$ y una subsucesión $\{x_{n_j}\} \subset \{x_n\}$ tales que $x_{n_j} \xrightarrow{\omega} x$.

Entonces, el funcional $\mathcal{J}_{\psi,\eta}$ definido en (1.12) tiene un minimizante global. Si además ψ es convexo y \mathcal{T} es inyectivo; o si ψ es estrictamente convexo, entonces dicho minimizante es único.

Observación 1.20. Si \mathcal{X} es un espacio de Banach reflexivo y $\psi(\cdot)$ es una norma definida sobre D entonces se satisfacen las hipótesis (H1) (tomando $\gamma = 0$), (H2) (por semicontinuidad inferior débil de una norma) y (H3) (por ser \mathcal{X} un espacio reflexivo) y por lo tanto, el funcional $\mathcal{J}_{\psi,\eta}$ definido en (1.12) tiene un minimizante global en D . Análogamente, si \mathcal{X} es un espacio de Banach reflexivo y ψ es el penalizante clásico dado por $\psi(x) = \|x\|^2$, es fácil ver que se satisfacen las hipótesis (H1), (H2) y (H3), de donde se sigue que el funcional \mathcal{J}_η dado en (1.8) tiene un minimizante global en D .

Observación 1.21. Notar que la existencia y la unicidad del minimizante global de $\mathcal{J}_{\psi,\eta}$ siguen siendo válidas si reemplazamos las hipótesis (H2) y (H3) sobre el funcional ψ por las siguientes hipótesis:

(H2'): ψ es ψ -ssi, es decir, para cada sucesión ψ -acotada $\{x_n\} \subset D$ tal que $x_n \rightarrow x \in D$, existe una subsucesión $\{x_{n_j}\} \subset \{x_n\}$ tal que $\psi(x) \leq \liminf_{j \rightarrow \infty} \psi(x_{n_j})$;

(H3'): las sucesiones ψ -acotadas son relativamente compactas en D , es decir, para cada sucesión ψ -acotada $\{x_n\} \subset D$, existen $x \in D$ y una subsucesión $\{x_{n_j}\} \subset \{x_n\}$ tales que $x_{n_j} \rightarrow x$.

Es oportuno notar que la hipótesis (H3') es más fuerte que (H3) mientras que (H2') es más débil que (H2). Veamos primero que (H3') implica (H3). En efecto, sea $\{x_n\} \subset D$ una sucesión ψ -acotada. Luego, por (H3') existen $x \in D$ y una subsucesión $\{x_{n_j}\} \subset \{x_n\}$ tales que $x_{n_j} \rightarrow x$. Entonces, $x_{n_j} \xrightarrow{\omega} x$ y así se tiene (H3). Ahora, veamos que (H2) implica (H2'). En efecto, sea $\{x_n\} \subset D$ una sucesión ψ -acotada tal que $x_n \rightarrow x \in D$. Luego, se tiene que $x_n \xrightarrow{\omega} x$ entonces por (H2) existe una subsucesión $\{x_{n_j}\} \subset \{x_n\}$ tal que $\psi(x) \leq \liminf_{j \rightarrow \infty} \psi(x_{n_j})$, de donde resulta (H2').

Si bien las hipótesis (H1), (H2) y (H3), así como (H2') y (H3'), imponen condiciones sobre el penalizante ψ y no sobre el operador \mathcal{T} , es posible relajarlas si se dispone de cierta información sobre \mathcal{T} en relación a ψ , como se presenta en el siguiente resultado.

Teorema 1.22. Sean \mathcal{X} un espacio vectorial normado, \mathcal{Y} un espacio con producto interno, $\mathcal{T} \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$, $y \in \mathcal{Y}$, $D \subset \mathcal{X}$ un conjunto convexo y ψ un funcional real sobre D . Consideremos además que \mathcal{T} y ψ satisfacen las siguientes hipótesis:

(H1): ψ está acotado por abajo, es decir, existe $\gamma \geq 0$ tal que $\psi(x) \geq -\gamma$, para todo $x \in D$;

(I2): para cada sucesión $\{x_n\} \subset D$ tal que $\{\|\mathcal{T}x_n\| + \psi(x_n)\}$ está acotada en \mathbb{R} (decimos en este caso que tal sucesión es \mathcal{T} - ψ -acotada) y $x_n \xrightarrow{\omega} x \in D$, existe una subsucesión $\{x_{n_j}\} \subset \{x_n\}$ tal que $\psi(x) \leq \liminf_{j \rightarrow \infty} \psi(x_{n_j})$;

(I3): las sucesiones \mathcal{T} - ψ -acotadas son relativamente débilmente compactas en D , es decir, para cada sucesión \mathcal{T} - ψ -acotada $\{x_n\} \subset D$, existen $x \in D$ y una subsucesión $\{x_{n_j}\} \subset \{x_n\}$ tales que $x_{n_j} \xrightarrow{\omega} x$.

Entonces el funcional $\mathcal{J}_{\psi, \eta}$ definido en (1.12) tiene un minimizante global. Si además, ψ es convexo y si \mathcal{T} es inyectivo; o si ψ es estrictamente convexo, entonces dicho minimizante es único.

Observación 1.23. Las hipótesis (I2) e (I3) son más débiles que (H2) y (H3), respectivamente. En primer lugar, veamos que (H2) implica (I2). En efecto, sea $\{x_n\} \subset D$ una sucesión \mathcal{T} - ψ -acotada tal que $x_n \xrightarrow{\omega} x \in D$. Luego, existe $k_0 \in \mathbb{N}$ tal que

$$\|\|\mathcal{T}x_n\| + \psi(x_n)\| \leq k_0, \quad \forall n \in \mathbb{N}. \quad (1.13)$$

Supongamos que $\{x_n\}$ no es ψ -acotada. Entonces, existe una subsucesión $\{x_{n_j}\} \subset \{x_n\}$ tal que $\psi(x_{n_j}) \rightarrow +\infty$ (como consecuencia también de (H1)). Así, existe $j_0 \in \mathbb{N}$ tal que $\psi(x_{n_j}) > k_0 > 0$, $\forall j \geq j_0$, de donde se sigue para $j \geq j_0$ que

$$\|\|\mathcal{T}x_{n_j}\| + \psi(x_{n_j})\| \geq \|\mathcal{T}x_{n_j}\| + \psi(x_{n_j}) \geq \psi(x_{n_j}) > k_0,$$

lo cual contradice (1.13). Por lo tanto, $\{x_n\}$ es ψ -acotada. Entonces, por (H2) existe una subsucesión $\{x_{n_j}\} \subset \{x_n\}$ tal que $\psi(x) \leq \liminf_{j \rightarrow \infty} \psi(x_{n_j})$, de donde se obtiene (I2).

Veamos ahora que (H3) implica (I3). Para ello, sea $\{x_n\} \subset D$ una sucesión \mathcal{T} - ψ -acotada. Análogamente a lo anterior, se tiene que $\{x_n\}$ es ψ -acotada. Por (H3), existen $x \in D$ y una subsucesión $\{x_{n_j}\} \subset \{x_n\}$ tales que $x_{n_j} \xrightarrow{\omega} x$, de donde se sigue (I3).

Observación 1.24. Notemos que el Teorema 1.22 sigue siendo válido si se reemplazan las hipótesis (I2) e (I3) por las siguientes:

(I2'): ψ es \mathcal{T} - ψ -ssi, es decir, para cada sucesión \mathcal{T} - ψ -acotada $\{x_n\} \subset D$ tal que $x_n \rightarrow x \in D$, existe una subsucesión $\{x_{n_j}\} \subset \{x_n\}$ tal que $\psi(x) \leq \liminf_{j \rightarrow \infty} \psi(x_{n_j})$;

(I3'): para cada sucesión \mathcal{T} - ψ -acotada $\{x_n\} \subset D$, existen $x \in D$ y una subsucesión $\{x_{n_j}\} \subset \{x_n\}$ tales que $x_{n_j} \rightarrow x$.

1.3. Regularización de tipo Tikhonov-Phillips doblemente generalizado

En la sección anterior presentamos el concepto de regularización de un problema inverso en el contexto de espacios de Banach y propusimos un método que consistió en aproximar la solución de este problema mediante un minimizante del funcional TPG, es decir, un funcional de Tikhonov-Phillips con penalizante generalizado. Sin embargo, en muchas aplicaciones, resulta de gran utilidad el uso de otros términos de fidelidad diferentes del clásico $\|\mathcal{T}x - y\|^2$. Por ejemplo, la utilización de normas más débiles como término de fidelidad permite, en el caso de problemas de restauración de imágenes, una separación “más adecuada” entre la componente de alta frecuencia, tales como el ruido y la textura, y la parte suave de la imagen ([41], [42]).

En el Capítulo 2, donde estudiaremos los problemas inversos desde el punto de vista estocástico, veremos que el término de fidelidad se encuentra estrechamente relacionado a la distribución de probabilidad de una variable aleatoria que representa un error aditivo en el modelo. De esta manera, bajo dicho enfoque, es usual pensar en el uso de términos de fidelidad diferentes del clásico. A continuación, presentamos dos fidelidades no cuadráticas.

Ejemplo 1.25. Como es bien sabido, el método de regularización por variación acotada introducido por L. I. Rudin, S. Osher y E. Fatemi en 1992, es muy apropiado en problemas de restauración de imágenes y eliminación de ruido cuando la solución exacta presenta discontinuidades y bordes. Sin embargo, este método en su versión original (donde el término de fidelidad es el clásico) tiene ciertas limitaciones. Una de ellas, es la pérdida de contraste en las soluciones incluso para imágenes sin ruido ([51]). En 2005, T. F. Chan y S. Esedoglu ([11]) estudiaron una versión modificada del método de variación acotada que tiene como término de fidelidad la norma L^1 (es decir, $\|\mathcal{T}x - y\|_1$) y la aplicaron a problemas de eliminación de ruido. Sus resultados permitieron afirmar que esa modificación produjo diferencias cualitativas interesantes y muy deseables en el comportamiento del método. En particular, debido a la forma en que se escalan el término de fidelidad y el penalizante en el método modificado, mostraron que el mismo, a diferencia del método original, es invariante por contraste. Además, la regularización impuesta a las soluciones del método con la norma L^1 es “más geométrica” en el sentido que el proceso de regularización depende más de las formas que del contraste de las características de la imagen. En el Capítulo 2 veremos que esta fidelidad se encuentra asociada a un error aditivo de tipo impulso con distribución de Laplace de media cero.

Ejemplo 1.26. En aplicaciones tales como Tomografía por Emisión de Positrones y Nanoscopía Óptica, los datos muestreados usualmente responden a un Proceso de Poisson. Por esta razón, como veremos en el Capítulo 2, el término de fidelidad más utilizado en este contexto es la divergencia de Kullback-Leibler $\text{KL}(\mathcal{T}x, y)$, donde KL es el funcional de Kullback-Leibler dado por

$$\text{KL}(u, v) \doteq \int_{\Omega} \left(v \ln \left(\frac{v}{u} \right) - v + u \right) d\mu, \quad (1.14)$$

donde $\Omega \subset \mathbb{R}^n$ es un conjunto acotado y medible, $u, v \in L^1(\Omega)$, $u, v > 0$ en c.t.p ([45]). En este contexto, el operador \mathcal{T} asociado al problema es lineal y compacto y satisface que $\mathcal{T}x > 0$ en c.t.p, para $x > 0$ en c.t.p.

En esta sección estudiaremos el método que consiste en aproximar la solución del problema inverso (1.1) mediante un minimizante de un funcional de Tikhonov-Phillips que sea generalizado tanto en el penalizante como en el término de fidelidad. Consideraremos entonces funcionales TPGG dados por

$$\mathcal{J}_{\phi, \psi, \eta}(x) \doteq \phi(\mathcal{T}x, y) + \eta \psi(x), \quad (1.15)$$

donde $\phi : \mathcal{Y} \times \mathcal{Y} \rightarrow \mathbb{R}$ y $\psi : D \subset \mathcal{X} \rightarrow \mathbb{R}$ son funcionales arbitrarios y η es una constante real positiva. Como hemos mencionado en la sección anterior, para determinar si este método propuesto constituye una regularización en el sentido de la Definición 1.14, es necesario probar la existencia y unicidad de soluciones regularizadas, la existencia de solución ψ -minimizante y, por último, la convergencia en algún sentido de las soluciones regularizadas a esta solución cuando el nivel de ruido tiende a cero.

A continuación, presentamos un resultado que asegura la existencia y unicidad de un minimizante del funcional (1.15). Este resultado constituye una generalización del Teorema 1.19.

Teorema 1.27. Sean \mathcal{X} e \mathcal{Y} espacios normados, $\mathcal{T} \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$, $y \in \mathcal{Y}$ y $D \subset \mathcal{X}$ un conjunto convexo. Supongamos que los funcionales $\psi : D \rightarrow \mathbb{R}$ y $\phi : \mathcal{Y} \times \mathcal{Y} \rightarrow \mathbb{R}$ satisfacen las siguientes hipótesis:

(G1): existen constantes reales $\gamma_1, \gamma_2 \geq 0$ tales que

- $\psi(x) \geq -\gamma_1$, para todo $x \in D$;
- $\phi(\bar{y}, \hat{y}) \geq -\gamma_2$, para todo $\bar{y}, \hat{y} \in \mathcal{Y}$;

(H2): para cada sucesión ψ -acotada $\{x_n\} \subset D$ tal que $x_n \xrightarrow{\omega} x \in D$, existe una subsucesión $\{x_{n_j}\} \subset \{x_n\}$ que verifica $\psi(x) \leq \liminf_{j \rightarrow \infty} \psi(x_{n_j})$;

(H3): las sucesiones ψ -acotadas son relativamente débilmente compactas en D , es decir, para cada sucesión ψ -acotada $\{x_n\} \subset D$, existen $x \in D$ y una subsucesión $\{x_{n_j}\} \subset \{x_n\}$ tales que $x_{n_j} \xrightarrow{\omega} x$;

(G4): para cada sucesión $\{y_k\} \subset \mathcal{Y}$ tal que $y_k \xrightarrow{\omega} \bar{y}$ se satisface $\phi(\bar{y}, \hat{y}) \leq \liminf_{k \rightarrow \infty} \phi(y_k, \hat{y})$, para todo $\hat{y} \in \mathcal{Y}$.

Entonces, el funcional $\mathcal{J}_{\phi, \psi, \eta}$ definido en (1.15) tiene un minimizante global. Si además, $\phi(\cdot, y)$ es estrictamente convexo, ψ es convexo y \mathcal{T} es inyectivo; o si $\phi(\cdot, y)$ es convexo y ψ es estrictamente convexo, entonces dicho minimizante es único.

Demostración. En primer lugar, notemos que para cada sucesión $\{z_n\} \subset D$ que converge débilmente a $z \in \mathcal{X}$, como \mathcal{T} es lineal y acotado se tiene que $\mathcal{T}z_n \xrightarrow{\omega} \mathcal{T}z$. Luego, por hipótesis (G4) para cada $y \in \mathcal{Y}$ se tiene que

$$\phi(\mathcal{T}z, y) \leq \liminf_{n \rightarrow \infty} \phi(\mathcal{T}z_n, y). \quad (1.16)$$

Dado que el funcional \mathcal{J} resulta acotado inferiormente (como consecuencia de (G1)), existe una sucesión $\{x_n\} \subset D$ tal que

$$\mathcal{J}_{\phi, \psi, \eta}(x_n) \longrightarrow \inf_{x \in D} \mathcal{J}_{\phi, \psi, \eta}(x) \doteq \mathcal{J}_{\min}. \quad (1.17)$$

Por hipótesis (G1) y como $\eta > 0$ es inmediato que $-\infty < \mathcal{J}_{\min}$. Supongamos ahora que $\mathcal{J}_{\min} = +\infty$. Como consecuencia, $\mathcal{J}_{\phi, \psi, \eta}(x_n) = +\infty$ para todo n , y así se prueba la existencia de un minimizante. Consideremos entonces el caso en que $\mathcal{J}_{\min} < +\infty$. Veamos ahora que $\mathcal{J}_{\phi, \psi, \eta}$ es un funcional ψ -coercivo. En efecto, sea $\{z_n\} \subset D$ tal que $\lim_{n \rightarrow \infty} \psi(z_n) = +\infty$. Luego, como $\eta > 0$ y ϕ satisface la hipótesis (G1) resulta que

$$\mathcal{J}_{\phi, \psi, \eta}(z_n) = \phi(\mathcal{T}z_n, y) + \eta \psi(z_n) \geq -\gamma_2 + \eta \psi(z_n) \longrightarrow +\infty.$$

Por lo tanto, $\lim_{n \rightarrow \infty} \mathcal{J}_{\phi, \psi, \eta}(z_n) = +\infty$ y así $\mathcal{J}_{\phi, \psi, \eta}$ es un funcional ψ -coercivo.

Supongamos ahora que la sucesión $\{x_n\}$ no es ψ -acotada. Entonces, existe una subsucesión $\{x_{n_j}\} \subset \{x_n\}$ tal que $\psi(x_{n_j}) \rightarrow +\infty$. Luego, como el funcional $\mathcal{J}_{\phi, \psi, \eta}$ es ψ -coercivo, $\mathcal{J}_{\phi, \psi, \eta}(x_{n_j}) \rightarrow +\infty$. Pero esto, junto a (1.17), contradice el hecho de que $\mathcal{J}_{\min} < +\infty$. Por lo tanto, $\{x_n\}$ es ψ -acotada. Entonces, por hipótesis (H3) existen $\bar{x} \in D$ y una subsucesión $\{x_{n_j}\} \subset \{x_n\}$ tales que $x_{n_j} \xrightarrow{\omega} \bar{x}$. Además, por hipótesis (H2) existe una subsucesión $\{x_{n_{j_k}}\} \subset \{x_{n_j}\}$ tal que

$$\psi(\bar{x}) \leq \liminf_{k \rightarrow \infty} \psi(x_{n_{j_k}}). \quad (1.18)$$

Luego,

$$\begin{aligned}
\mathcal{J}_{\phi,\psi,\eta}(\bar{x}) &= \phi(\mathcal{T}\bar{x}, y) + \eta\psi(\bar{x}) \\
&\leq \liminf_{k \rightarrow \infty} \phi(\mathcal{T}x_{n_{j_k}}, y) + \eta\psi(\bar{x}) && \text{(por (1.16))} \\
&\leq \liminf_{k \rightarrow \infty} \phi(\mathcal{T}x_{n_{j_k}}, y) + \eta \liminf_{k \rightarrow \infty} \psi(x_{n_{j_k}}) && \text{(por (1.18) y } \eta > 0) \\
&\leq \liminf_{k \rightarrow \infty} \left(\phi(\mathcal{T}x_{n_{j_k}}, y) + \eta\psi(x_{n_{j_k}}) \right) && \text{(por propiedad de } \liminf \text{ y } \eta > 0) \\
&= \liminf_{k \rightarrow \infty} \mathcal{J}_{\phi,\psi,\eta}(x_{n_{j_k}}) \\
&= \mathcal{J}_{\min}. && \text{(por (1.17) puesto que } \{x_{n_{j_k}}\} \subset \{x_n\})
\end{aligned}$$

Dado que $\mathcal{J}_{\min} \leq \mathcal{J}_{\phi,\psi,\eta}(\bar{x})$ (por definición de ínfimo), se tiene que $\mathcal{J}_{\min} = \mathcal{J}_{\phi,\psi,\eta}(\bar{x})$, lo cual prueba la existencia de un minimizante global del funcional (1.15).

Para probar ahora la unicidad del minimizante es necesario, en primer lugar, demostrar que el funcional $\mathcal{J}_{\phi,\psi,\eta}$ es estrictamente convexo. Consideraremos dos casos. En primer lugar, supongamos que $\phi(\cdot, y)$ es estrictamente convexo, ψ es convexo y \mathcal{T} es inyectivo. Sean $x, z \in D$, $x \neq z$ y $\beta \in (0, 1)$. Como \mathcal{T} es inyectivo se tiene que $\mathcal{T}x \neq \mathcal{T}z$ y así,

$$\begin{aligned}
\mathcal{J}_{\phi,\psi,\eta}(\beta x + (1 - \beta)z) &= \phi(\mathcal{T}(\beta x + (1 - \beta)z), y) + \eta\psi(\beta x + (1 - \beta)z) \\
&< \beta\phi(\mathcal{T}x, y) + (1 - \beta)\phi(\mathcal{T}z, y) + \eta\psi(\beta x + (1 - \beta)z) && (\phi(\cdot, y) \text{ estrict. convexo y } \mathcal{T} \text{ lineal}) \\
&\leq \beta\phi(\mathcal{T}x, y) + (1 - \beta)\phi(\mathcal{T}z, y) + \eta\beta\psi(x) + \eta(1 - \beta)\psi(z) && (\psi \text{ convexo}) \\
&= \beta\mathcal{J}_{\phi,\psi,\eta}(x) + (1 - \beta)\mathcal{J}_{\phi,\psi,\eta}(z).
\end{aligned}$$

Entonces, el funcional $\mathcal{J}_{\phi,\psi,\eta}$ es estrictamente convexo. Veamos ahora el segundo caso en donde suponemos que $\phi(\cdot, y)$ es convexo y ψ es estrictamente convexo. Sean $x, z \in D$, $x \neq z$ y $\beta \in (0, 1)$. Luego,

$$\begin{aligned}
\mathcal{J}_{\phi,\psi,\eta}(\beta x + (1 - \beta)z) &= \phi(\mathcal{T}(\beta x + (1 - \beta)z), y) + \eta\psi(\beta x + (1 - \beta)z) \\
&< \phi(\beta\mathcal{T}x + (1 - \beta)\mathcal{T}z, y) + \eta\beta\psi(x) + \eta(1 - \beta)\psi(z) && (\psi \text{ estrict. convexo y } \mathcal{T} \text{ lineal}) \\
&\leq \beta\phi(\mathcal{T}x, y) + (1 - \beta)\phi(\mathcal{T}z, y) + \eta\beta\psi(x) + \eta(1 - \beta)\psi(z) && (\phi(\cdot, y) \text{ convexo}) \\
&= \beta\mathcal{J}_{\phi,\psi,\eta}(x) + (1 - \beta)\mathcal{J}_{\phi,\psi,\eta}(z),
\end{aligned}$$

de donde se sigue que el funcional $\mathcal{J}_{\phi,\psi,\eta}$ es estrictamente convexo.

Supongamos ahora que existen $x_1, x_2 \in D$, $x_1 \neq x_2$ tales que $\mathcal{J}_{\phi,\psi,\eta}(x_1) = \mathcal{J}_{\phi,\psi,\eta}(x_2) = \mathcal{J}_{\min}$. Como D es un conjunto convexo, para $\beta \in (0, 1)$ sabemos que $\beta x_1 + (1 - \beta)x_2 \in D$. Además,

como $\mathcal{J}_{\phi,\psi,\eta}$ es estrictamente convexo se tiene que

$$\begin{aligned} \mathcal{J}_{\phi,\psi,\eta}(\beta x_1 + (1 - \beta)x_2) &< \beta \mathcal{J}_{\phi,\psi,\eta}(x_1) + (1 - \beta) \mathcal{J}_{\phi,\psi,\eta}(x_2) \\ &= \mathcal{J}_{\phi,\psi,\eta}(x_1). \end{aligned} \quad (\text{pues } \mathcal{J}_{\phi,\psi,\eta}(x_1) = \mathcal{J}_{\phi,\psi,\eta}(x_2))$$

Pero esto contradice el hecho que $\mathcal{J}_{\phi,\psi,\eta}(x_1) = \min_{x \in \mathcal{D}} \mathcal{J}_{\phi,\psi,\eta}(x)$. Por lo tanto, $x_1 = x_2$, de donde se obtiene la unicidad del minimizante global. \square

Si consideramos el término de fidelidad cuadrático $\phi(\mathcal{T}x, y) \doteq \|\mathcal{T}x - y\|^2$, es sencillo ver que el mismo verifica la hipótesis (G1) con $\gamma_2 = 0$ y la hipótesis (G4) por la semicontinuidad inferior débil de la norma. De esta manera, el Teorema 1.27 constituye una generalización del Teorema 1.19 para funcionales de Tikhonov-Phillips con término de fidelidad y penalizante generalizados. Sin embargo, es posible considerar términos de fidelidad diferentes del clásico que verifican las hipótesis (G1) y (G4). En particular, las fidelidades presentadas en los ejemplos 1.25 y 1.26 son algunas de ellas. En efecto, es fácil probar (análogamente a lo realizado para el caso clásico) que la fidelidad dada por $\phi(\mathcal{T}x, y) \doteq \|\mathcal{T}x - y\|_1$ verifica las hipótesis (G1) y (G4) del teorema anterior por tratarse de una norma. Por otro lado, es sabido que el funcional de Kullback-Leibler es no negativo ([15]) por lo que verifica la hipótesis (G1) con $\gamma_2 = 0$. Además, dicho funcional es débilmente semicontinuo inferiormente en la primera variable ([15], Corolario 2.2), con lo cual verifica la hipótesis (G4).

Presentamos ahora una generalización del Teorema 1.22 que relaja las hipótesis (H2) y (H3) del Teorema 1.27 utilizando cierta información sobre \mathcal{T} en relación a ϕ y a ψ .

Teorema 1.28. Sean \mathcal{X} e \mathcal{Y} espacios normados, $\mathcal{T} \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$, $y \in \mathcal{Y}$, $D \subset \mathcal{X}$ un conjunto convexo, y $\psi : D \rightarrow \mathbb{R}$, $\phi : \mathcal{Y} \times \mathcal{Y} \rightarrow \mathbb{R}$ funcionales que satisfacen las siguientes hipótesis:

(G1): existen constantes reales $\gamma_1, \gamma_2 \geq 0$ tales que

- $\psi(x) \geq -\gamma_1$, para todo $x \in D$;
- $\phi(\bar{y}, \hat{y}) \geq -\gamma_2$, para todo $\bar{y}, \hat{y} \in \mathcal{Y}$;

(G2): para cada sucesión $\{x_n\} \subset D$ tal que $\{\phi(\mathcal{T}x_n, y) + \psi(x_n)\}$ está acotada en \mathbb{R} (decimos en este caso que tal sucesión es ϕ - \mathcal{T} - ψ -acotada) y $x_n \xrightarrow{\omega} x \in D$, existe una subsucesión $\{x_{n_j}\} \subset \{x_n\}$ tal que $\psi(x) \leq \liminf_{j \rightarrow \infty} \psi(x_{n_j})$;

(G3): las sucesiones ϕ - \mathcal{T} - ψ -acotadas son relativamente débilmente compactas en D , es decir, para cada sucesión ϕ - \mathcal{T} - ψ -acotada $\{x_n\} \subset D$, existen $x \in D$ y una subsucesión $\{x_{n_j}\} \subset \{x_n\}$ tales que $x_{n_j} \xrightarrow{\omega} x$.

(G4): para cada sucesión $\{y_k\} \subset \mathcal{Y}$ tal que $y_k \xrightarrow{\omega} \bar{y}$ se satisface $\phi(\bar{y}, \hat{y}) \leq \liminf_{k \rightarrow \infty} \phi(y_k, \hat{y})$, para todo $\hat{y} \in \mathcal{Y}$.

Entonces el funcional $\mathcal{J}_{\phi, \psi, \eta}$ definido en (1.15) tiene un minimizante global. Si además, $\phi(\cdot, y)$ es estrictamente convexo, ψ es convexo y \mathcal{T} es inyectivo; o si $\phi(\cdot, y)$ es convexo y ψ es estrictamente convexo, entonces dicho minimizante es único.

Demostración. Si $\{z_n\} \subset D$ es una sucesión tal que $z_n \xrightarrow{\omega} z \in \mathcal{X}$ entonces se puede probar, como lo hicimos en la demostración del Teorema 1.27, que

$$\phi(\mathcal{T}z, y) \leq \liminf_{n \rightarrow \infty} \phi(\mathcal{T}z_n, y). \quad (1.19)$$

Análogamente a la prueba del Teorema 1.27, sabemos que existe una sucesión $\{x_n\} \subset D$ tal que

$$\mathcal{J}_{\phi, \psi, \eta}(x_n) \longrightarrow \inf_{x \in D} \mathcal{J}_{\phi, \psi, \eta}(x) \doteq \mathcal{J}_{\min}. \quad (1.20)$$

y que $-\infty < \mathcal{J}_{\min} < +\infty$. Supongamos que la sucesión $\{x_n\}$ no es ϕ - \mathcal{T} - ψ -acotada. Entonces, existe una subsucesión $\{x_{n_j}\} \subset \{x_n\}$ tal que $\phi(\mathcal{T}x_{n_j}, y) + \psi(x_{n_j}) \rightarrow +\infty$. Consideremos los siguientes casos:

- Si $\phi(\mathcal{T}x_{n_j}, y) \rightarrow +\infty$ entonces por hipótesis (G1) para ψ se sigue que

$$\mathcal{J}_{\phi, \psi, \eta}(x_{n_j}) = \phi(\mathcal{T}x_{n_j}, y) + \eta \psi(x_{n_j}) \geq \phi(\mathcal{T}x_{n_j}, y) - \eta \gamma_1 \longrightarrow +\infty,$$

y así $\mathcal{J}_{\phi, \psi, \eta}(x_{n_j}) \rightarrow +\infty$. Luego, por unicidad del límite, $\mathcal{J}_{\min} = +\infty$, lo que contradice el hecho que $\mathcal{J}_{\min} < +\infty$.

- Si $\psi(x_{n_j}) \rightarrow +\infty$ entonces por hipótesis (G1) para ϕ y como $\eta > 0$ se sigue que

$$\mathcal{J}_{\phi, \psi, \eta}(x_{n_j}) = \phi(\mathcal{T}x_{n_j}, y) + \eta \psi(x_{n_j}) \geq -\gamma_2 + \eta \psi(x_{n_j}) \longrightarrow +\infty,$$

de donde resulta que $\mathcal{J}_{\phi, \psi, \eta}(x_{n_j}) \rightarrow +\infty$. Así, como en el caso anterior, se contradice que $\mathcal{J}_{\min} < +\infty$.

Por lo tanto, la sucesión $\{x_n\}$ es ϕ - \mathcal{T} - ψ -acotada y entonces, por hipótesis (G3) existen $\bar{x} \in D$ y una subsucesión $\{x_{n_j}\} \subset \{x_n\}$ tales que $x_{n_j} \xrightarrow{\omega} \bar{x}$. Luego, por (G2) existe una subsucesión $\{x_{n_{j_k}}\} \subset \{x_{n_j}\}$ tal que

$$\psi(\bar{x}) \leq \liminf_{k \rightarrow \infty} \psi(x_{n_{j_k}}). \quad (1.21)$$

Entonces,

$$\begin{aligned} \mathcal{J}_{\phi, \psi, \eta}(\bar{x}) &= \phi(\mathcal{T}\bar{x}, y) + \eta \psi(\bar{x}) \\ &\leq \liminf_{k \rightarrow \infty} \phi(\mathcal{T}x_{n_{j_k}}, y) + \eta \psi(\bar{x}) \end{aligned} \quad (\text{por (1.19)})$$

$$\begin{aligned}
&\leq \liminf_{k \rightarrow \infty} \phi(\mathcal{T}x_{n_{j_k}}, y) + \eta \liminf_{k \rightarrow \infty} \psi(x_{n_{j_k}}) && \text{(por (1.21) y } \eta > 0) \\
&\leq \liminf_{k \rightarrow \infty} \left(\phi(\mathcal{T}x_{n_{j_k}}, y) + \eta \psi(x_{n_{j_k}}) \right) && \text{(por propiedad de límite inferior y } \eta > 0) \\
&= \liminf_{k \rightarrow \infty} \mathcal{J}_{\phi, \psi, \eta}(x_{n_{j_k}}) \\
&= \mathcal{J}_{\min}. && \text{(por (1.20) puesto que } \{x_{n_{j_k}}\} \subset \{x_n\})
\end{aligned}$$

Por lo tanto, usando la definición de ínfimo se sigue que $\mathcal{J}_{\min} = \mathcal{J}_{\phi, \psi, \eta}(\bar{x})$ y así, existe un minimizante global del funcional $\mathcal{J}_{\phi, \psi, \eta}$.

Análogamente a la prueba del Teorema 1.27, si $\phi(\cdot, y)$ es estrictamente convexo, ψ es convexo y \mathcal{T} es inyectivo, o si $\phi(\cdot, y)$ es convexo y ψ es estrictamente convexo, entonces el funcional $\mathcal{J}_{\phi, \psi, \eta}$ resulta estrictamente convexo y, finalmente, como D es un conjunto convexo, se prueba la unicidad del minimizante. \square

Observación 1.29. Análogamente a lo realizado en la Observación 1.23, puede mostrarse que las hipótesis (G2) y (G3) son más débiles que las hipótesis (H2) y (H3), respectivamente. En primer lugar, para ver que (H2) implica (G2), consideremos una sucesión ϕ - \mathcal{T} - ψ -acotada $\{x_n\} \subset D$ tal que $x_n \xrightarrow{\omega} x \in D$. Análogamente a la Observación 1.23, es fácil ver que esta sucesión resulta ψ -acotada. Entonces, por (H2) existe una subsucesión $\{x_{n_j}\} \subset \{x_n\}$ tal que $\psi(x) \leq \liminf_{j \rightarrow \infty} \psi(x_{n_j})$, de donde se sigue inmediatamente (G2).

Veamos ahora que (H3) implica (G3). Para ello, sea $\{x_n\} \subset D$ una sucesión ϕ - \mathcal{T} - ψ -acotada. Luego, se tiene que $\{x_n\}$ es ψ -acotada. Por (H3), existen $x \in D$ y una subsucesión $\{x_{n_j}\} \subset \{x_n\}$ tales que $x_{n_j} \xrightarrow{\omega} x$, de donde se sigue (G3).

Hasta aquí hemos probado la existencia y la unicidad de soluciones regularizadas obtenidas mediante la minimización de un funcional TPGG. Procedemos ahora a probar la existencia de una solución ψ -minimizante.

Teorema 1.30. Sean \mathcal{X} e \mathcal{Y} espacios normados, $y \in \mathcal{Y}$ y $\mathcal{T} \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$. Supongamos que el funcional $\psi : D \subset \mathcal{X} \rightarrow \mathbb{R}$ satisface las siguientes hipótesis:

(H1): ψ está acotado por abajo, es decir, existe $\gamma \geq 0$ tal que $\psi(x) \geq -\gamma$, para todo $x \in D$;

(H2): para cada sucesión ψ -acotada $\{x_n\} \subset D$ tal que $x_n \xrightarrow{\omega} x \in D$, existe una subsucesión $\{x_{n_j}\} \subset \{x_n\}$ que verifica $\psi(x) \leq \liminf_{j \rightarrow \infty} \psi(x_{n_j})$;

(H3): las sucesiones ψ -acotadas son relativamente débilmente compactas en D , es decir, para cada sucesión ψ -acotada $\{x_k\} \subset D$, existen $x \in D$ y una subsucesión $\{x_{k_j}\} \subset \{x_k\}$ tales que $x_{k_j} \xrightarrow{\omega} x$.

Si existe una solución de (1.1) entonces existe una solución ψ -minimizante de (1.1) en D .

Demostración. Sea $c \doteq \inf_{x \in D: \mathcal{T}x=y} \{\psi(x)\}$ (notar que existe como consecuencia de (H1)). Luego, existe una sucesión $\{x_k\}$ de soluciones de (1.1) en D tal que $\psi(x_k) \rightarrow c$. Dado que la sucesión $\{\psi(x_k)\}$ es convergente en \mathbb{R} , se sigue que $\{x_k\}$ es una sucesión ψ -acotada. Luego, por hipótesis (H3) existe una subsucesión $\{x_{k_j}\} \subset \{x_k\}$ y $\tilde{x} \in D$ tal que $x_{k_j} \xrightarrow{\omega} \tilde{x}$. A su vez, por hipótesis (H2), existe una subsucesión $\{x_{k_{j_n}}\}$ de $\{x_{k_j}\}$ tal que

$$\psi(\tilde{x}) \leq \liminf_{n \rightarrow \infty} \psi(x_{k_{j_n}}) = c. \quad (1.22)$$

Como \mathcal{T} es un operador lineal y acotado se tiene que $\mathcal{T}x_{k_{j_n}} \xrightarrow{\omega} \mathcal{T}\tilde{x}$. Luego, del hecho de que $\mathcal{T}x_{k_{j_n}} = y$, para todo n se sigue inmediatamente que $\mathcal{T}\tilde{x} = y$. Por lo tanto, como consecuencia de la desigualdad (1.22) y debido a que $\tilde{x} \in D$ se tiene que $\psi(\tilde{x}) = c$ y, así, \tilde{x} es una solución ψ -minimizante de (1.1) en D . \square

Observación 1.31. Cabe mencionar aquí que, tal como era de esperarse, para garantizar la existencia de una solución ψ -minimizante del problema (1.1) se requieren sólo hipótesis sobre el penalizante y no así sobre el término de fidelidad.

Finalmente, con el propósito de probar que los métodos propuestos constituyen una regularización en el sentido de la Definición 1.14, sólo resta probar la convergencia de las soluciones regularizadas a una solución ψ -minimizante del problema (1.1). Esto lo haremos para el caso en que el parámetro de regularización sea elegido mediante una regla a-priori. Para ello, dada una sucesión de datos $\{y^{\delta_k}\} \subset \mathcal{Y}$ tales que $\phi(y, y^{\delta_k}) \leq \delta_k$ y $\delta_k \rightarrow 0^+$ cuando $k \rightarrow \infty$, debemos probar que la sucesión de soluciones regularizadas $\{x_{\eta(\delta_k)}^{\delta_k}\} \subset \mathcal{X}$ obtenidas mediante la minimización del funcional TPGG converge, en algún sentido, a una solución ψ -minimizante del problema (1.1).

Teorema 1.32. Sean \mathcal{X} e \mathcal{Y} espacios normados, $y \in \mathcal{Y}$ y $\mathcal{T} \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$. Supongamos que los funcionales $\psi : D \subset \mathcal{X} \rightarrow [0, +\infty)$ y $\phi : \mathcal{Y} \times \mathcal{Y} \rightarrow [0, +\infty]$ satisfacen las siguientes hipótesis:

- I. para cada sucesión ψ -acotada $\{x_n\} \subset D$ tal que $x_n \xrightarrow{\omega} x \in D$, existe una subsucesión $\{x_{n_j}\} \subset \{x_n\}$ que verifica $\psi(x) \leq \liminf_{j \rightarrow \infty} \psi(x_{n_j})$;
- II. las sucesiones ψ -acotadas son relativamente débilmente compactas en D , es decir, para cada sucesión ψ -acotada $\{x_n\} \subset D$, existen $x \in D$ y una subsucesión $\{x_{n_j}\} \subset \{x_n\}$ tales que $x_{n_j} \xrightarrow{\omega} x$;
- III. si $\phi(y, y_n) \rightarrow 0^+$ entonces $y_n \xrightarrow{\omega} y$;

- IV. si $y_n \xrightarrow{\omega} y$ y $\bar{y}_n \xrightarrow{\omega} \bar{y}$ entonces $\phi(y, \bar{y}) \leq \liminf_{n \rightarrow \infty} \phi(y_n, \bar{y}_n)$;
 V. $\phi(y, \bar{y}) = 0$ es equivalente a $y = \bar{y}$.

Supongamos además que existe una solución de (1.1) y que

- $\{\delta_k\}$ es una sucesión decreciente de niveles de ruido tales que $\delta_k \rightarrow 0^+$;
- $\{y^{\delta_k}\} \subset \mathcal{Y}$ es una sucesión de datos tales que $\phi(y, y^{\delta_k}) \leq \delta_k$ para todo k ;
- $\eta(\delta)$ es una regla de elección del parámetro a-priori tal que

$$\eta(\delta) \rightarrow 0 \quad y \quad \frac{\delta}{\eta(\delta)} \rightarrow 0 \quad \text{cuando } \delta \rightarrow 0^+.$$

Consideremos la sucesión $\{x_{\eta(\delta_k)}^{\delta_k}\}$ de soluciones regularizadas que son minimizantes del funcional $\mathcal{J}_{\phi, \psi, \eta(\delta_k)}$ con dato y^{δ_k} . Entonces, toda subsucesión de $\{x_{\eta(\delta_k)}^{\delta_k}\}$ contiene una subsucesión débilmente convergente a una solución ψ -minimizante de (1.1). Si además, la solución ψ -minimizante es única entonces $\{x_{\eta(\delta_k)}^{\delta_k}\}$ converge débilmente a dicha solución.

Demostración. Con el objetivo de simplificar la notación, denotemos con $x_k \doteq x_{\eta(\delta_k)}^{\delta_k}$, $y_k \doteq y^{\delta_k}$ y $\eta_k \doteq \eta(\delta_k)$.

Sean $\{x_{k_j}\}_j$ una subsucesión de $\{x_k\}_k$ y \tilde{x} una solución ψ -minimizante de (1.1) (\tilde{x} existe como consecuencia del Teorema 1.30). Dado que x_{k_j} es minimizante del funcional $\mathcal{J}_{\phi, \psi, \eta_{k_j}}$ para el dato y_{k_j} , los funcionales ψ y ϕ son no negativos y por ser \tilde{x} una solución ψ -minimizante se sigue que

$$\begin{aligned} 0 \leq \eta_{k_j} \psi(x_{k_j}) &\leq \phi(\mathcal{T}x_{k_j}, y_{k_j}) + \eta_{k_j} \psi(x_{k_j}) \\ &\leq \phi(\mathcal{T}\tilde{x}, y_{k_j}) + \eta_{k_j} \psi(\tilde{x}) \\ &= \phi(y, y_{k_j}) + \eta_{k_j} \psi(\tilde{x}) \\ &\leq \delta_{k_j} + \eta_{k_j} \psi(\tilde{x}), \end{aligned}$$

y así, $0 \leq \psi(x_{k_j}) \leq \frac{\delta_{k_j}}{\eta_{k_j}} + \psi(\tilde{x})$. Por lo tanto, como $\frac{\delta_{k_j}}{\eta_{k_j}} \rightarrow 0^+$ y $\psi(\tilde{x}) < +\infty$, se tiene que

$$\limsup_{j \rightarrow \infty} \psi(x_{k_j}) \leq \psi(\tilde{x}). \quad (1.23)$$

Entonces la sucesión $\{x_{k_j}\}$ resulta ser ψ -acotada y, por hipótesis II, existen una subsucesión $\{x_{k_{j_n}}\}_n$ de $\{x_{k_j}\}_j$ y $\hat{x} \in D$ tales que $x_{k_{j_n}} \xrightarrow{\omega} \hat{x}$. Luego, por ser \mathcal{T} un operador lineal y acotado se sigue que

$$\mathcal{T}x_{k_{j_n}} \xrightarrow{\omega} \mathcal{T}\hat{x}. \quad (1.24)$$

Por otro lado, por ser $x_{k_{j_n}}$ minimizante del funcional $\mathcal{J}_{\phi, \psi, \eta_{k_{j_n}}}$ para el dato $y_{k_{j_n}}$, \tilde{x} una solución ψ -minimizante y debido a que los funcionales ψ y ϕ son no negativos se tiene que

$$\begin{aligned} 0 \leq \phi(\mathcal{T}x_{k_{j_n}}, y_{k_{j_n}}) &\leq \phi(\mathcal{T}x_{k_{j_n}}, y_{k_{j_n}}) + \eta_{k_{j_n}} \psi(x_{k_{j_n}}) \leq \phi(\mathcal{T}\tilde{x}, y_{k_{j_n}}) + \eta_{k_{j_n}} \psi(\tilde{x}) \\ &\leq \delta_{k_{j_n}} + \eta_{k_{j_n}} \psi(\tilde{x}). \end{aligned}$$

Luego, como $\delta_{k_{j_n}} \rightarrow 0^+$, $\eta_{k_{j_n}} \rightarrow 0^+$ y $\psi(\tilde{x}) < +\infty$ se tiene que

$$\lim_{n \rightarrow \infty} \phi(\mathcal{T}x_{k_{j_n}}, y_{k_{j_n}}) = 0. \quad (1.25)$$

Ahora, por hipótesis 0 $0 \leq \phi(y, y_{k_{j_n}}) \leq \delta_{k_{j_n}}$ y $\delta_{k_{j_n}} \rightarrow 0^+$, de donde se sigue inmediatamente por hipótesis III que

$$y_{k_{j_n}} \xrightarrow{\omega} y. \quad (1.26)$$

Entonces, de (1.24), (1.25), (1.26) y por hipótesis IV resulta

$$0 \leq \phi(\mathcal{T}\hat{x}, y) \leq \liminf_{n \rightarrow \infty} \phi(\mathcal{T}x_{k_{j_n}}, y_{k_{j_n}}) = 0,$$

de donde se sigue que $\phi(\mathcal{T}\hat{x}, y) = 0$. Por lo tanto, como consecuencia de la hipótesis v, $\mathcal{T}\hat{x} = y$ y así \hat{x} es una solución de (1.1). Veamos ahora que \hat{x} es una solución ψ -minimizante. En efecto, como $\{x_{k_{j_n}}\}$ es una sucesión ψ -acotada y $x_{k_{j_n}} \xrightarrow{\omega} \hat{x}$, por hipótesis I existe una subsucesión, que volvemos a denotar de la misma manera para simplificar la notación, tal que $\psi(\hat{x}) \leq \liminf_{n \rightarrow \infty} \psi(x_{k_{j_n}})$. Luego,

$$\begin{aligned} \psi(\hat{x}) &\leq \liminf_{n \rightarrow \infty} \psi(x_{k_{j_n}}) \\ &\leq \limsup_{n \rightarrow \infty} \psi(x_{k_{j_n}}) \\ &\leq \psi(\tilde{x}) && \text{(por (1.23))} \\ &\leq \psi(\hat{x}), && \text{(por ser } \tilde{x} \text{ una solución } \psi\text{-minimizante)} \end{aligned} \quad (1.27)$$

de donde se sigue que $\psi(\hat{x}) = \psi(\tilde{x})$, es decir que \hat{x} es una solución ψ -minimizante de (1.1).

Finalmente, si \tilde{x} es única, hemos probado que para toda subsucesión $\{x_{k_j}\}_j$ de $\{x_k\}_k$ existe una subsucesión $\{x_{k_{j_n}}\}_n$ que converge débilmente a \tilde{x} . Por lo tanto, $x_k \xrightarrow{\omega} \tilde{x}$. \square

Observación 1.33. Cabe mencionar aquí que las hipótesis I y II del teorema anterior coinciden con las hipótesis (H2) y (H3) del Teorema 1.27 y, tal como hemos mencionado en la Observación 1.29, las mismas implican las hipótesis (G2) y (G3) del Teorema 1.28. Por lo tanto, si reemplazamos las hipótesis I y II del teorema anterior por (G2) y (G3), bajo el supuesto de existencia de solución ψ -minimizante, se prueba la convergencia de soluciones regularizadas.

Observación 1.34. Observar que de la desigualdad (1.27) y del hecho que $\psi(\hat{x}) = \psi(\tilde{x})$ se tiene que $\lim_{n \rightarrow \infty} \psi(x_{k_{j_n}}) = \psi(\hat{x}) = \psi(\tilde{x})$. Así, probamos que para toda subsucesión $\{x_{k_j}\}_j$ de $\{x_k\}_k$ existe una subsucesión $\{x_{k_{j_n}}\}_n$ tal que $\psi(x_{k_{j_n}})$ converge a $\psi(\tilde{x})$. Por lo tanto, $\lim_{k \rightarrow \infty} \psi(x_k) = \psi(\tilde{x})$. Supongamos además que ψ es un funcional tal que verifica la siguiente implicación: *para cada sucesión $\{x_n\} \subset \mathcal{X}$ tal que $x_n \xrightarrow{\omega} x \in \mathcal{X}$ y $\psi(x_n) \rightarrow \psi(x)$ se tiene que $\|x_n - x\| \rightarrow 0$* . Luego, bajo las hipótesis del Teorema 1.32 y la unicidad de \tilde{x} , se prueba la convergencia fuerte de soluciones regularizadas a la solución ψ -minimizante \tilde{x} .

El siguiente resultado es consecuencia del Teorema 1.32 y asegura la convergencia de soluciones regularizadas obtenidas mediante la minimización de un funcional TPG. Su prueba se sigue inmediatamente verificando que el término de fidelidad cuadrático $\phi(\mathcal{T}x, y) \doteq \|\mathcal{T}x - y\|^2$ cumple las hipótesis III, IV y V del Teorema 1.32.

Corolario 1.35. Sean \mathcal{X} e \mathcal{Y} espacios normados, $y \in \mathcal{Y}$ y $\mathcal{T} \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$. Supongamos que el funcional $\psi : D \subset \mathcal{X} \rightarrow [0, +\infty)$ satisface las siguientes hipótesis:

- I. para cada sucesión ψ -acotada $\{x_n\} \subset D$ tal que $x_n \xrightarrow{\omega} x \in D$, existe una subsucesión $\{x_{n_j}\} \subset \{x_n\}$ que verifica $\psi(x) \leq \liminf_{j \rightarrow \infty} \psi(x_{n_j})$;
- II. las sucesiones ψ -acotadas son relativamente débilmente compactas en D , es decir, para cada sucesión ψ -acotada $\{x_n\} \subset D$, existen $x \in D$ y una subsucesión $\{x_{n_j}\} \subset \{x_n\}$ tales que $x_{n_j} \xrightarrow{\omega} x$.

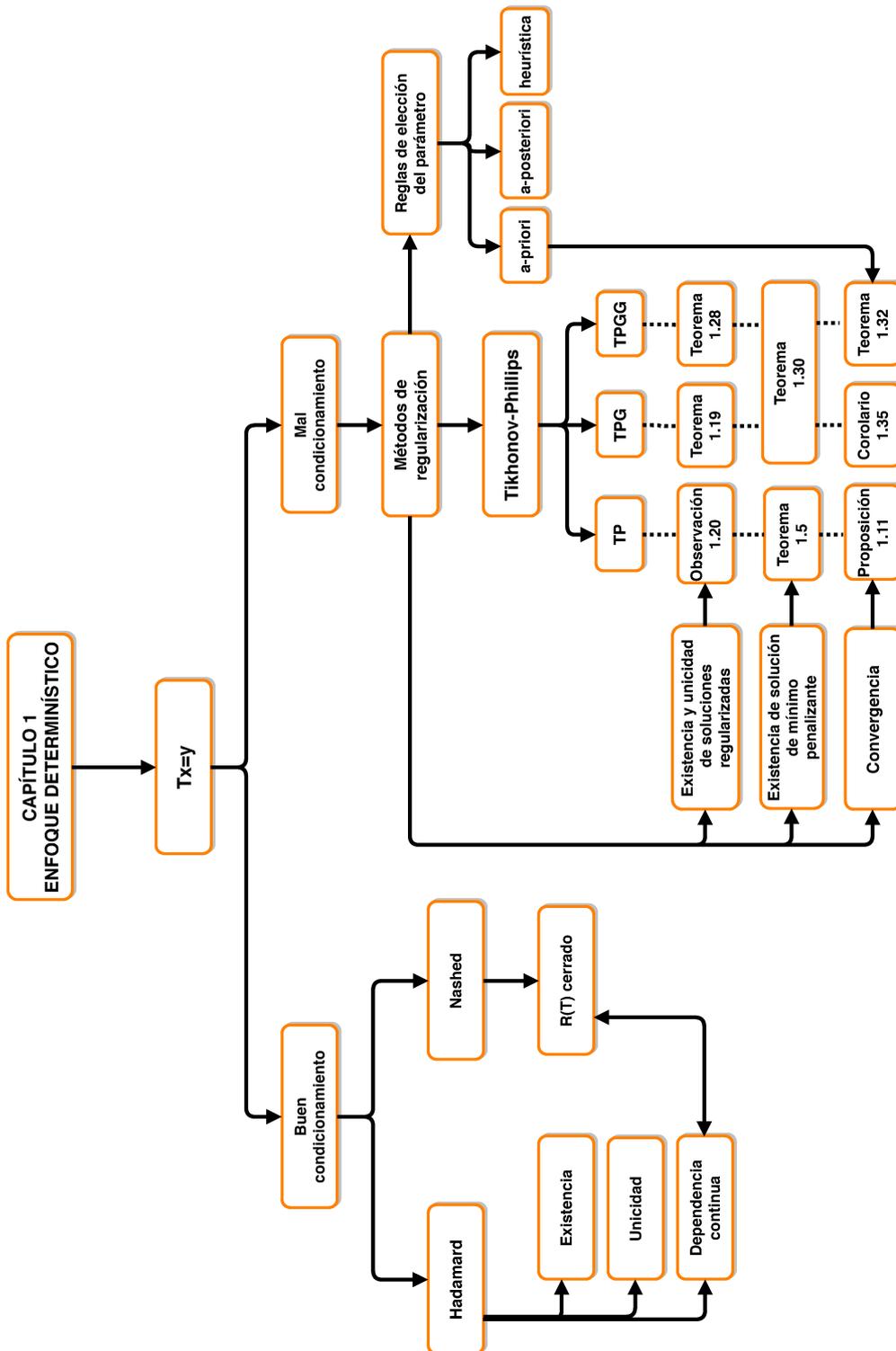
Supongamos además que existe una solución de (1.1) y que

- $\{\delta_k\}$ es una sucesión decreciente de niveles de ruido tales que $\delta_k \rightarrow 0^+$;
- $\{y^{\delta_k}\} \subset \mathcal{Y}$ es una sucesión de datos tales que $\|y - y^{\delta_k}\|^2 \leq \delta_k$, para todo k ;
- $\eta(\delta)$ es una regla de elección del parámetro a-priori tal que

$$\eta(\delta) \rightarrow 0 \quad \text{y} \quad \frac{\delta}{\eta(\delta)} \rightarrow 0 \quad \text{cuando} \quad \delta \rightarrow 0^+.$$

Consideremos la sucesión $\{x_{\eta(\delta_k)}^{\delta_k}\}$ de soluciones regularizadas que son minimizantes del funcional $\mathcal{J}_{\psi, \eta(\delta_k)}$ con dato y^{δ_k} . Entonces, toda subsucesión de $\{x_{\eta(\delta_k)}^{\delta_k}\}$ contiene una subsucesión débilmente convergente a una solución ψ -minimizante de (1.1). Si además, la solución ψ -minimizante es única entonces $\{x_{\eta(\delta_k)}^{\delta_k}\}$ converge débilmente a dicha solución.

En conclusión, de lo expuesto hasta aquí en las Secciones 1.2 y 1.3 se sigue que los métodos propuestos, que consisten en aproximar la solución del problema (1.1) mediante un minimizante de los funcionales TPG o TPGG, constituyen regularizaciones en el sentido de la Definición 1.14.



CAPÍTULO 2

Problemas inversos desde el enfoque estocástico

Existen, en principio, dos grandes enfoques para resolver problemas inversos: uno clásico determinístico, el cual presentamos en el Capítulo 1, y otro estocástico, que veremos en el presente capítulo. Si bien, en una primera instancia, ambos enfoques parecen disjuntos, a lo largo del desarrollo de este capítulo mostraremos que se encuentran estrechamente relacionados.

Supongamos que se desea resolver el problema inverso que consiste en obtener información acerca de una variable de interés $x \in \mathbb{R}^n$ a partir de una variable observable $y \in \mathbb{R}^m$, las cuales se relacionan a través de un modelo de la forma $y = f(x)$. Como hemos mencionado en el capítulo anterior, en la mayoría de los casos, el dato y no se conoce con exactitud puesto que puede contener errores (de medición, por ejemplo) y, de esta manera, un modelo adecuado para este tipo de problemas viene dado por

$$y = f(x, e), \tag{2.1}$$

donde $f : \mathbb{R}^n \times \mathbb{R}^k \rightarrow \mathbb{R}^m$ es una función asociada al modelo y el vector $e \in \mathbb{R}^k$ representa el error en la medición. En el Capítulo 1 hemos mencionado que, en el enfoque determinístico, el propósito de las técnicas clásicas de regularización es obtener una aproximación *razonable* de la variable de interés basada en los datos de los que se dispone. Sin embargo, desde el punto de vista estadístico, la solución de un problema inverso no será una estimación puntual de la variable incógnita sino una distribución de probabilidad que podrá utilizarse para obtener, luego, estimaciones de la variable de interés.

La filosofía detrás de este nuevo enfoque consiste en la *reformulación* del problema inverso en una forma estadística de búsqueda de información, es decir, en un problema de inferencia. Así, todas las variables son consideradas variables aleatorias absolutamente continuas de manera tal que la distribución de probabilidad de cada una de ellas puede ser expresada en términos de una función de densidad de probabilidad. A lo largo de este capítulo utilizaremos, como es usual, letras mayúsculas para representar las variables aleatorias y letras minúsculas para sus realizaciones. De esta manera, el modelo (2.1) puede reescribirse como

$$Y = f(X, E). \tag{2.2}$$

Supongamos ahora que, antes de obtener la medición y , disponemos de cierta información a-priori sobre la variable incógnita X y que la misma puede codificarse en una densidad de probabilidad π_{pr} , que llamamos *densidad a-priori*. Por otro lado, supongamos que luego de la medición puede obtenerse la densidad de probabilidad conjunta de X e Y que denotamos con $\pi(x, y)$. Luego, la densidad de probabilidad de Y dado $X = x$ con $\pi_{\text{pr}}(x) > 0$, llamada *función de verosimilitud*, viene dada por

$$\pi(y|x) = \frac{\pi(x, y)}{\pi_{\text{pr}}(x)}.$$

Finalmente, dada una observación y de Y con $\pi(y) > 0$, la densidad condicional

$$\pi(x|y) = \frac{\pi(x, y)}{\pi(y)}$$

se denomina *distribución a-posteriori* de X y se considera como la solución de un problema inverso desde este enfoque. Una forma de determinar la distribución a-posteriori es utilizando el Teorema de Bayes que enunciamos a continuación.

Teorema 2.1. *Sea X un vector aleatorio en \mathbb{R}^n con función de densidad de probabilidad a-priori dada por π_{pr} . Sea y una observación del vector aleatorio $Y \in \mathbb{R}^m$ tal que $\pi(y) > 0$. Entonces, la distribución de probabilidad a-posteriori de X dado el dato y está dada por*

$$\pi_{\text{post}}(x) \doteq \pi(x|y) = \frac{\pi_{\text{pr}}(x) \pi(y|x)}{\pi(y)} \propto \pi_{\text{pr}}(x) \pi(y|x),$$

donde el símbolo \propto significa “proporcional a”.

Por lo tanto, desde el punto de vista *Bayesiano*, un problema inverso puede expresarse de la siguiente manera: *dada una observación y de la variable Y , hallar la distribución de probabilidad condicional $\pi(x|y)$ de la variable incógnita X ([29]).* Claramente, una distribución de probabilidades como “solución” a un problema inverso resulta poco útil. Es por ello que, una vez obtenida dicha distribución, a partir de ella se hallan estimadores de la variable de interés. En consecuencia, la resolución de un problema inverso desde el punto de vista estadístico puede resumirse en la realización de las siguientes tres tareas:

- Tarea 1. Construir una densidad de probabilidad π_{pr} que refleje el comportamiento esperado de variable de interés X basado en la información a-priori que se dispone de dicha variable.
- Tarea 2. Encontrar la función de verosimilitud $\pi(y|x)$ que describe la relación entre la observación y la incógnita.
- Tarea 3. Hallar la densidad a-posteriori π_{post} , e implementar métodos para explorarla.

A lo largo de la Sección 2.1 describiremos distintas formas de llevar a cabo estas tres tareas ([29]). En la Sección 2.1.1 presentaremos algunas densidades a-priori que serán de especial interés en este capítulo. En la Sección 2.1.2 veremos cómo construir la función de verosimilitud para el caso particular de ruido aditivo y presentaremos el proceso de Poisson como ejemplo de modelo con ruido no aditivo. Luego, en la Sección 2.1.3 definiremos, a partir de la distribución a-posteriori, algunos estimadores de la variable incógnita. En la Sección 2.2 analizaremos el principal vínculo que existe entre los enfoques determinístico y estocástico para el abordaje de un problema inverso. Este análisis permitirá ver que hallar un estimador de la variable de interés X , está fuertemente conectado con la minimización de un funcional $\text{TPG}(G)$. A continuación, en la Sección 2.3 abordaremos los modelos jerárquicos ([9], [29]) y, finalmente, presentaremos una regla de elección del parámetro heurística cuya construcción se basa en este tipo de modelo ([27]). Por último, en la Sección 2.4 mostraremos algunos ejemplos numéricos en aplicaciones de procesamiento de señales e imágenes con el principal objetivo de mostrar el comportamiento de la regla presentada.

2.1. Tareas

2.1.1. Densidad a-priori

Como hemos mencionado, la primer tarea que debe realizarse para la resolución de un problema inverso desde el enfoque Bayesiano es la construcción de una densidad a-priori para la variable incógnita X . Esta tarea es crucial para la resolución y, en muchos casos, constituye un gran desafío. La mayor dificultad se encuentra en lograr transformar la información cualitativa que disponemos de la variable incógnita en información cuantitativa y, de esta manera, encontrar una densidad a-priori adecuada que refleje el comportamiento esperado de X . Por ejemplo, en imágenes médicas, podríamos estar interesados en determinar la presencia de cierto tumor del que se sabe su ubicación exacta y que posiblemente tiene cierta estructura superficial característica, que el ojo entrenado de un radiólogo puede reconocer. Estas son descripciones cualitativas de lo que esperaríamos observar en la imagen, pero difíciles de traducir al lenguaje de las densidades. El principal objetivo entonces es diseñar una densidad a-priori π_{pr} , de manera tal que si E es un conjunto de realizaciones esperadas del vector aleatorio X y U es un conjunto de realizaciones no esperadas, se verifique que $\pi_{\text{pr}}(x) \gg \pi_{\text{pr}}(\tilde{x})$, para $x \in E$ y $\tilde{x} \in U$. De esta manera, la densidad a-priori debe concentrarse en esos valores de la variable X que esperamos ver y asignarles una probabilidad claramente mayor que a aquellos que no esperamos ver.

En esta sección, a través de ejemplos, presentaremos algunas densidades conocidas que serán utilizadas a lo largo de este capítulo, tanto como densidades a-priori para la variable de interés como densidades para el ruido, haciendo especial hincapié en algunas densidades de impulso (ℓ^1 , Gamma y Gamma Inversa), en las densidades Gaussianas y en la distribución de Poisson. Además, con el objetivo de visualizar el comportamiento de la variable según la distribución elegida, se muestrean dichas densidades.

Ejemplo 2.2. Una variable aleatoria $X \in \mathbb{R}$ tiene una *distribución de Laplace* de media μ y varianza $2\sigma^2$ ($\sigma > 0$) si su densidad de probabilidad está dada por

$$\pi(x) = \frac{1}{2\sigma} \exp\left(-\frac{|x - \mu|}{\sigma}\right).$$

Supongamos que $X \in \mathbb{R}^n$ es una imagen (vectorizada) cuya componente X_j es una variable aleatoria que representa la intensidad en el j -ésimo píxel. Supongamos además que dichas componentes son independientes e idénticamente distribuídas y tienen una distribución de Laplace de media μ y varianza $2\sigma^2$. Luego, la densidad de probabilidad del vector aleatorio X está dada por

$$\pi(x) = \frac{1}{(2\sigma)^n} \exp\left(-\sum_{j=1}^n \frac{|x_j - \mu|}{\sigma}\right).$$

Si $\mu = 0$ entonces decimos que X tiene una distribución dada por la *densidad de impulso* ℓ^1 definida como

$$\pi(x) = \left(\frac{\alpha}{2}\right)^n \exp(-\alpha\|x\|_1), \quad (2.3)$$

donde $\alpha \doteq \frac{1}{\sigma} > 0$ es un parámetro conocido y $\|x\|_1 \doteq \sum_{j=1}^n |x_j|$. Denotamos esto con $X \sim \ell^1(\alpha)$.

Como X representa una imagen, es natural considerar una restricción de positividad para la misma, resultando necesario redefinir la densidad ℓ^1 como

$$\pi(x) = \alpha^n \pi_+(x) \exp(-\alpha\|x\|_1), \quad (2.4)$$

donde $\pi_+(x) = 1$, si $x_j > 0$, $\forall j$ y $\pi_+(x) = 0$, en otro caso. Cabe destacar aquí que esta nueva función cumple con las condiciones necesarias para ser una densidad.

Es importante observar aquí que la densidad (2.4) corresponde a la distribución conjunta de n variables aleatorias independientes X_j . Luego, para cada píxel se puede tomar una muestra de tamaño 1 utilizando el procedimiento usual conocido como *Método de la Transformada Inversa* ([46]). En la Figura 1 se muestra una imagen de 60×60 píxeles obtenida a partir de este procedimiento considerando $\alpha = 10$.

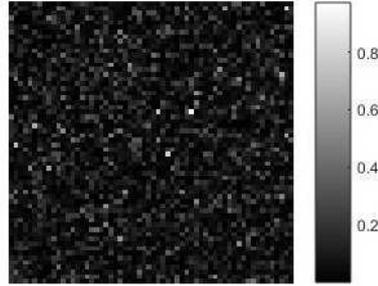


FIGURA 1. Muestra de una densidad $\ell^1(10)$ con restricción de positividad.

Ejemplo 2.3. Existen otras densidades de impulso como las *distribuciones Gamma* y *Gamma Inversa* de parámetros α y θ_0 . Para un vector aleatorio $X \in \mathbb{R}^n$ con componentes independientes, estas densidades son de la forma

$$\pi_G(x) \propto \prod_{j=1}^n x_j^{\alpha-1} \exp\left(-\frac{x_j}{\theta_0}\right),$$

y

$$\pi_{GI}(x) \propto \prod_{j=1}^n x_j^{-\alpha-1} \exp\left(-\frac{\theta_0}{x_j}\right),$$

respectivamente. Denotamos estos casos con $X \sim \Gamma(\alpha, \theta_0)$ y $X \sim \Gamma^{-1}(\alpha, \theta_0)$.

Es sabido que estas distribuciones, al igual que la densidad ℓ^1 , favorecen cambios repentinos en la variable, permitiendo la presencia de datos atípicos (outliers). La principal diferencia entre estas dos distribuciones es que la distribución Gamma Inversa favorece soluciones con menor cantidad de datos atípicos pero con magnitudes mayores que aquellos obtenidos con la distribución Gamma. Esto puede observarse en la Figura 2.

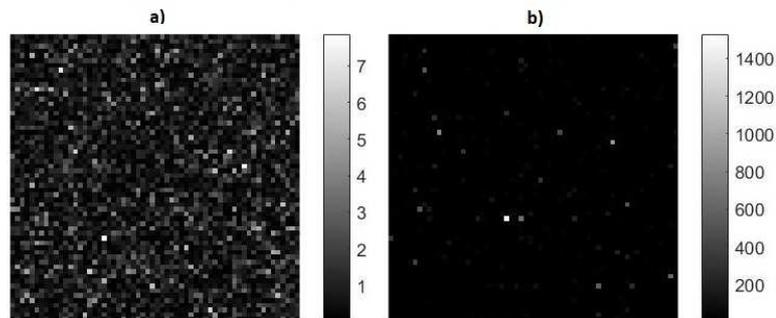


FIGURA 2. Muestra de una densidad: a) $\Gamma(1, 1)$; b) $\Gamma^{-1}(1, 1)$.

En la Figura 2 se presentan dos imágenes de 60×60 píxeles obtenidas a partir de una distribución Gamma (parte a)) y de una distribución Gamma Inversa (parte b)) ambas con parámetros $\alpha = 1$ y $\theta_0 = 1$.

Ejemplo 2.4. Las distribuciones Gaussianas son de gran importancia en el área estadística puesto que permiten modelizar numerosos fenómenos naturales, sociales y psicológicos.

Sean $x_0 \in \mathbb{R}^n$ y $\Gamma \in \mathbb{R}^{n \times n}$ una matriz simétrica definida positiva. Una variable aleatoria n -dimensional *Gaussiana* X con media x_0 y matriz de covarianza Γ es una variable aleatoria con densidad de probabilidad dada por

$$\pi(x) = \left(\frac{1}{(2\pi)^n |\Gamma|} \right)^{1/2} \exp \left(-\frac{1}{2} (x - x_0)^t \Gamma^{-1} (x - x_0) \right).$$

Denotamos esto con $X \sim \mathcal{N}(x_0, \Gamma)$.

Consideremos una matriz de covarianza Γ asociada a una discretización del Laplaciano (por diferencias finitas con condiciones de borde Dirichlet). Esto permite construir una *densidad Gaussiana de suavidad* para el vector aleatorio X que tiene la forma

$$\pi(x) \propto \exp \left(-\frac{1}{2\gamma^2} \|Lx\|^2 \right),$$

donde γ es un parámetro positivo y, en principio, conocido. Denotamos esto con $X \sim \mathcal{N}_S(\gamma, L)$ (es decir, $X \sim \mathcal{N}(0, \gamma^2 (L^t L)^{-1})$). En la Figura 3 se muestra una imagen de 60×60 píxeles obtenida a partir de una distribución Gaussiana de suavidad con $\gamma = \frac{1}{2}$. Allí puede observarse el comportamiento “suave” esperado de la variable.

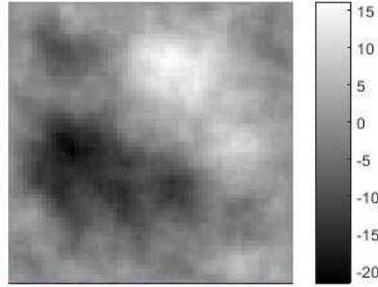


FIGURA 3. Muestra de una densidad $\mathcal{N}_S(1/2, L)$.

Ejemplo 2.5. Dado $x \in \mathbb{N}_0^n$, decimos que el vector X tiene una *distribución de Poisson* si su densidad está dada por

$$\pi(x) = \prod_{i=1}^n \frac{e^{-\lambda_i} \lambda_i^{x_i}}{x_i!},$$

donde $\lambda_i > 0$ es un parámetro conocido y representa la tasa de ocurrencia del evento que se desea contabilizar en la i -ésima coordenada del vector X . Denotamos esto con $X \sim P(\lambda)$, donde λ es un vector cuya i -ésima coordenada es λ_i . En la Figura 4 se muestra una imagen de 60×60 píxeles obtenida a partir de una distribución de Poisson de parámetro $\lambda_i = 5$ para todo i .

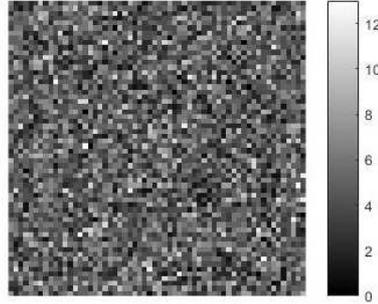


FIGURA 4. Muestra de una densidad de Poisson con $\lambda_i = 5 \forall i$.

Es oportuno mencionar que en cada uno de los muestreos realizados se utilizaron parámetros fijos. Si bien podría resultar interesante estudiar el comportamiento de las densidades frente a la variación de los mismos, no se modifica el comportamiento general esperado de la variable y por ello se omite este análisis en este trabajo.

Veamos ahora en qué casos se pueden utilizar las densidades presentadas anteriormente. Supongamos que se desea resolver un problema inverso en el cual la variable incógnita es una señal o imagen y que se sabe, además, que dicha imagen contiene pequeños objetos bien localizados, o para el caso de una señal, que contiene saltos o discontinuidades. Esto ocurre, por ejemplo, cuando uno está interesado en localizar un tumor en una imagen de rayos X. En estos casos es conveniente utilizar las densidades de impulso debido a que, como hemos visto, permiten la presencia de datos atípicos de grandes magnitudes que se distinguen de un fondo de amplitud baja. Por otro lado, si lo que se espera es que la imagen tenga un comportamiento “suave”, es decir que las intensidades de píxeles vecinos sean similares, se recomienda la utilización de una densidad Gaussiana de suavidad. Por otra parte, en muchas aplicaciones, las mediciones se basan en el recuento de ciertos eventos. Por ejemplo, podemos pensar en el caso de un microscopio electrónico en el que el dispositivo cuenta los electrones que llegan al detector durante un período de tiempo determinado. De manera similar, en imágenes de rayos X de baja energía, los datos consisten en conteos de fotones. En muchos casos, el número de conteos en cada observación puede ser muy bajo o simplemente cero. Una densidad que puede modelar este tipo de variables es la densidad de Poisson. De esta manera, dependiendo del tipo de comportamiento esperado de la variable incógnita, es posible elegir una densidad a-priori que represente dicho comportamiento.

2.1.2. Función de verosimilitud

La segunda tarea para la resolución de un problema inverso, que consiste en la construcción de la función de verosimilitud, suele ser frecuentemente el paso más directo de la teoría de inversión estadística. Es oportuno observar que la función de verosimilitud depende del tipo de relación que existe entre las variables (modelo aditivo, multiplicativo, etc) y, en muchas ocasiones, de la distribución asociada al ruido.

Consideremos, en primer lugar, que el ruido E es aditivo e independiente de la variable incógnita X , es decir, el modelo estocástico está dado por

$$Y = f(X) + E, \quad (2.5)$$

donde $X \in \mathbb{R}^n$, $Y, E \in \mathbb{R}^m$ y X y E son independientes. Supongamos además que la densidad de probabilidad del ruido es conocida y que está dada por π_E .

Si fijamos $X = x$ en (2.5), como X y E son independientes, la densidad de probabilidad de Y dado $X = x$ coincide con la densidad de probabilidad de E trasladada por $f(x)$, es decir

$$\pi(y|x) = \pi_E(y - f(x)).$$

Luego, si π_{pr} es la densidad de probabilidad a-priori para X , por el Teorema 2.1 tenemos que la densidad a-posteriori de X dado y es de la forma

$$\pi(x|y) \propto \pi_{\text{pr}}(x) \pi_E(y - f(x)).$$

A continuación, presentamos un ejemplo donde la variable E sigue una distribución Gaussiana y se considera un modelo con ruido aditivo e independiente de la variable de interés X .

Ejemplo 2.6. Supongamos que el ruido E es una variable aleatoria en \mathbb{R}^m con densidad Gaussiana de media cero y matriz de covarianzas $\sigma^2 I_m$, con σ conocido, la cual está dada por

$$\pi_E(e) = \frac{1}{(2\pi)^{m/2} \sigma^m} \exp\left(-\frac{1}{2\sigma^2} \|e\|^2\right).$$

En consecuencia, bajo el modelo con ruido aditivo (2.5), la función de verosimilitud resulta

$$\pi(y|x) = \frac{1}{(2\pi)^{m/2} \sigma^m} \exp\left(-\frac{1}{2\sigma^2} \|y - f(x)\|^2\right).$$

Existen otros tipos de ruido como, por ejemplo, ruido multiplicativo, que no serán tratados en esta tesis (para más información ver [29]).

En algunos casos la función de verosimilitud no se basa en un modelo del tipo $Y = f(X, E)$. En particular, veremos un ejemplo en donde se conoce la densidad de probabilidad de la variable Y pero la misma contiene un parámetro que depende de la variable incógnita X .

Ejemplo 2.7. Como hemos mencionado en la sección anterior, en muchas aplicaciones, las mediciones consisten en el recuento de ciertos eventos. En tales casos, la variable $Y \in \mathbb{N}_0^m$ puede modelarse como un *proceso de Poisson* el cual supone que cada componente Y_i es una variable aleatoria con distribución de Poisson y que su observación esperada \bar{y}_i viene dada por el modelo

$$\bar{y}_i = f_i(x),$$

donde $x \in \mathbb{R}^n$ es una observación de la variable incógnita X . Si además todas las componentes de Y son independientes entonces la función de verosimilitud resulta

$$\pi(y|x) = \prod_{i=1}^m \frac{\exp(-f_i(x)) f_i(x)^{y_i}}{y_i!}.$$

2.1.3. Densidad a-posteriori

La tercera y última tarea que debe realizarse para la resolución de un problema inverso desde el enfoque estocástico es la obtención y exploración de la densidad a-posteriori con el objetivo de estimar la variable incógnita. Una forma sencilla de obtener esta densidad es a partir del Teorema de Bayes. Es claro que la definición abstracta de la solución de un problema inverso como la distribución de probabilidad a-posteriori no es muy útil en la práctica. En general, se pretende hallar distintos estimadores a partir de dicha distribución. Entre los más utilizados se encuentran el estimador *máximo a-posteriori* y la *media condicional* que se definen a continuación.

Definición 2.8. Dada la densidad de probabilidad a-posteriori $\pi(x|y)$ del vector aleatorio $X \in \mathbb{R}^n$, se define:

- El estimador *máximo a-posteriori*, y se denota con x_{MAP} , como

$$x_{\text{MAP}} \doteq \arg \max_{x \in \mathbb{R}^n} \pi(x|y),$$

siempre que el máximo exista.

- El estimador *media condicional*, denotado con x_{MC} , como

$$x_{\text{MC}} \doteq E(x|y) = \int_{\mathbb{R}^n} x \pi(x|y) dx,$$

bajo el supuesto de que la integral exista.

En la próxima sección veremos que el estimador x_{MAP} , en ciertas ocasiones, se encuentra asociado a un minimizante de un funcional de tipo Tikhonov-Phillips generalizado o doblemente generalizado.

2.2. Vínculo entre los enfoques determinístico y estocástico

A continuación veremos el principal vínculo que existe entre los enfoques determinístico y estocástico para la resolución de un problema inverso mal condicionado. Bajo un modelo lineal con ruido aditivo, a partir de diferentes ejemplos, veremos cómo la densidad del error se relaciona con el término de fidelidad de un funcional de tipo Tikhonov-Phillips generalizado y cómo la densidad a-priori de la variable incógnita está estrechamente relacionada con el penalizante de dicho funcional. En los primeros ejemplos, consideraremos error con distribución Gaussiana y diferentes densidades a-priori para la variable de interés (Casos I, II, III y IV). Luego, presentaremos otros casos donde el error sigue una distribución no Gaussiana y, nuevamente, supondremos diferentes densidades a-priori para la incógnita (Casos V y VI). Finalmente, presentaremos un ejemplo de un modelo que no es de la forma $Y = f(X, E)$ y veremos nuevamente que la función de verosimilitud y la densidad a-priori de la variable incógnita se relacionan con el término de fidelidad y el penalizante de un funcional de tipo Tikhonov-Phillips doblemente generalizado, respectivamente (Caso VII). En todos los casos que estudiaremos, veremos que el estimador máximo a-posteriori asociado a la densidad a-posteriori de la variable de interés coincide con un minimizante de un funcional de tipo Tikhonov-Phillips generalizado o doblemente generalizado.

En los siguientes seis casos consideraremos un modelo lineal con ruido aditivo dado por

$$Y = TX + E,$$

donde $X \in \mathbb{R}^n$, $Y, E \in \mathbb{R}^m$ son variables aleatorias, X y E independientes y $T \in \mathbb{R}^{m \times n}$ es una matriz conocida asociada al modelo.

CASO I. $\mathcal{N} - \mathcal{N}$: Consideremos el caso donde $X \sim \mathcal{N}(0, \lambda^2 I_n)$ y supongamos que el ruido aditivo $E \sim \mathcal{N}(0, \tau^2 I_m)$, con τ conocido. Luego, del Teorema de Bayes para problemas inversos se sigue inmediatamente que la densidad a-posteriori de X dado $Y = y$ está dada por

$$\begin{aligned} \pi_{\text{post}}(x) &= \pi(x|y) \propto \pi_{\text{pr}}(x) \pi_e(y - Tx) \\ &\propto \exp\left(-\frac{1}{2\tau^2} \|y - Tx\|_2^2 - \frac{1}{2\lambda^2} \|x\|_2^2\right). \end{aligned}$$

De esta manera, el estimador máximo a-posteriori, x_{MAP} , de X estará dado por

$$\begin{aligned} x_{\text{MAP}} &= \arg \max_{x \in \mathbb{R}^n} \left\{ \exp \left(-\frac{1}{2\tau^2} \|y - Tx\|_2^2 - \frac{1}{2\lambda^2} \|x\|_2^2 \right) \right\} \\ &= \arg \max_{x \in \mathbb{R}^n} \left\{ \exp \left(-\frac{1}{2} \|y - Tx\|_2^2 - \frac{\eta}{2} \|x\|_2^2 \right) \right\} \\ &= \arg \min_{x \in \mathbb{R}^n} \left\{ \|y - Tx\|_2^2 + \eta \|x\|_2^2 \right\}, \end{aligned}$$

donde $\eta \doteq \frac{\tau^2}{\lambda^2}$. Es decir, el estimador x_{MAP} será un minimizante del funcional TP para el problema $Tx = y$ con parámetro de regularización η .

A partir de la teoría clásica de las distribuciones Gaussianas ([29]) es posible probar que la densidad a-posteriori de X dado $Y = y$ resulta Gaussiana con media (condicional) dada por

$$x_{\text{MC}} \doteq (T^t T + \eta I_n)^{-1} T^t y.$$

Luego, x_{MC} coincide con la solución regularizada (1.7) obtenida a partir de la minimización del funcional TP para el problema $Tx = y$, puesto que $T^* = T^t$ en este caso.

Finalmente, es oportuno destacar la relación existente entre el parámetro de regularización η y los parámetros σ y γ asociados a las densidades de E y X , respectivamente.

CASO II. $\mathcal{N} - \mathcal{N}_S$: Consideremos la densidad a-priori Gaussiana de suavidad de parámetro λ para X y supongamos el ruido aditivo $E \sim \mathcal{N}(0, \tau^2 I_m)$, con τ conocido. Luego, la densidad a-posteriori de X dado $Y = y$ resulta de la forma

$$\begin{aligned} \pi_{\text{post}}(x) &\propto \pi_{\text{pr}}(x) \pi_e(y - Tx) \\ &\propto \exp \left(-\frac{1}{2\tau^2} \|y - Tx\|_2^2 - \frac{1}{2\lambda^2} \|Lx\|_2^2 \right). \end{aligned}$$

De aquí se sigue inmediatamente que

$$\begin{aligned} x_{\text{MAP}} &= \arg \max_{x \in \mathbb{R}^n} \left\{ \exp \left(-\frac{1}{2} \|y - Tx\|_2^2 - \frac{\eta}{2} \|Lx\|_2^2 \right) \right\} \\ &= \arg \min_{x \in \mathbb{R}^n} \left\{ \|y - Tx\|_2^2 + \eta \|Lx\|_2^2 \right\}, \end{aligned}$$

donde $\eta \doteq \frac{\tau^2}{\lambda^2}$. De esta manera, el estimador x_{MAP} será, nuevamente, un minimizante de un funcional TPG con penalizante $\|Lx\|_2^2$ y parámetro de regularización η .

CASO III. $\mathcal{N} - \ell^1$: Consideremos ahora el caso donde $X \sim \ell^1(\lambda)$ y $E \sim \mathcal{N}(0, \tau^2 I_m)$, con τ conocido. Entonces, la densidad a-posteriori de X dado $Y = y$ es de la forma

$$\pi_{\text{post}}(x) \propto \exp\left(-\frac{1}{2\tau^2}\|y - Tx\|_2^2 - \lambda\|x\|_1\right).$$

Por lo tanto, el estimador máximo a-posteriori para X resulta

$$x_{\text{MAP}} = \arg \max_{x \in \mathbb{R}^n} \left\{ \exp\left(-\frac{1}{2\tau^2}\|y - Tx\|_2^2 - \lambda\|x\|_1\right) \right\} = \arg \min_{x \in \mathbb{R}^n} \{ \|y - Tx\|_2^2 + \eta\|x\|_1 \},$$

donde $\eta \doteq 2\tau^2\lambda$ y, así, x_{MAP} coincide con un minimizante de un funcional TPG con penalizante $\|x\|_1$ y parámetro de regularización η .

El siguiente caso tiene como objetivo generalizar los presentados anteriormente considerando una densidad a-priori más general para X . Esta densidad estará vinculada a la familia exponencial ([33]) que se define como el conjunto de las funciones de distribución tales que la función de densidad (o de probabilidad puntual) puede escribirse como

$$\pi(x; \theta) = \exp\left(\sum_{i=1}^k a_i(\theta) b_i(x) + c(\theta) + d(x)\right), \quad (2.6)$$

donde $\theta = (\theta_1, \dots, \theta_k)$ es un vector de parámetros asociados a la distribución de X y a_i, b_i, c, d son funciones reales.

CASO IV. $\mathcal{N} - \psi$: Consideremos ahora el caso donde X tiene una densidad a-priori dada por $\pi_{\text{pr}}(x; \lambda) = \exp(a(\lambda)\psi(x) + c(\lambda))$, donde el parámetro $\lambda > 0$ es conocido y $a(\lambda) < 0$ para todo λ . Cabe mencionar aquí que π_{pr} corresponde a una densidad asociada a la familia exponencial (2.6) considerando $d(x) = 0$. A su vez, tal como hemos mencionado π_{pr} generaliza las densidades consideradas en los casos anteriores puesto que $a(\lambda) = -\frac{1}{2\lambda^2}$, si $X \sim \mathcal{N}(0, \lambda^2 I_n)$ o $X \sim \mathcal{N}_S(\lambda, L)$, y $a(\lambda) = -\lambda$, si $X \sim \ell^1(\lambda)$, siendo en ambos casos una función negativa. Supongamos además $E \sim \mathcal{N}(0, \tau^2 I_m)$, con τ conocido. Entonces, se sigue inmediatamente que la densidad a-posteriori de X dado $Y = y$ es de la forma

$$\pi_{\text{post}}(x) \propto \exp\left(-\frac{1}{2\tau^2}\|y - Tx\|_2^2 + a(\lambda)\psi(x)\right).$$

Por lo tanto, el estimador máximo a-posteriori para X resulta

$$x_{\text{MAP}} = \arg \max_{x \in \mathbb{R}^n} \left\{ \exp\left(-\frac{1}{2\tau^2}\|y - Tx\|_2^2 + a(\lambda)\psi(x)\right) \right\} = \arg \min_{x \in \mathbb{R}^n} \{ \|y - Tx\|_2^2 + \eta\psi(x) \},$$

donde $\eta \doteq 2\tau^2(-a(\lambda))$ y, así, x_{MAP} coincide con un minimizante de un funcional TPG con penalizante $\psi(x)$ y parámetro de regularización η .

CASO V. $\ell^1 - \ell^1$: Consideremos ahora el caso donde tanto la variable de interés X como el ruido E tienen densidades a-priori de impulso ℓ^1 de parámetros conocidos λ y τ , respectivamente. Luego, la densidad a-posteriori de X dado $Y = y$ en este caso resulta

$$\pi_{\text{post}}(x) \propto \exp(-\tau\|y - Tx\|_1 - \lambda\|x\|_1).$$

Así, el estimador máximo a-posteriori para X resulta

$$x_{\text{MAP}} = \arg \max_{x \in \mathbb{R}^n} \{\exp(-\tau\|y - Tx\|_1 - \lambda\|x\|_1)\} = \arg \min_{x \in \mathbb{R}^n} \{\|y - Tx\|_1 + \eta\|x\|_1\},$$

donde $\eta \doteq \lambda/\tau$ y, así, x_{MAP} coincide con un minimizante de un funcional TPGG con fidelidad $\|y - Tx\|_1$, penalizante $\|x\|_1$ y parámetro de regularización η .

El siguiente caso generaliza los casos anteriores considerando que tanto la variable de interés X como el ruido E tienen distribuciones asociadas a la familia exponencial (2.6). Cabe mencionar aquí que las distribuciones consideradas en todos los casos anteriores pertenecen a esta familia con $d(x) = 0$ en cada caso.

CASO VI. $\phi - \psi$: Consideremos ahora el caso general donde X tiene una densidad a-priori dada por $\pi_{\text{pr}}(x; \lambda) = \exp(a_1(\lambda)\psi(x) + c_1(\lambda))$, con parámetro λ conocido y que el ruido E tiene una distribución dada por $\pi_E(e; \tau) = \exp(a_2(\tau)\phi(e) + c_2(\tau))$, donde el parámetro τ es conocido. Además, supongamos que $a_1(\lambda)$ y $a_2(\tau)$ tienen el mismo signo para todo λ y τ . Luego, la densidad a-posteriori de X dado $Y = y$ está dada por

$$\pi_{\text{post}}(x) \propto \exp(a_2(\tau)\phi(y - Tx) + a_1(\lambda)\psi(x)),$$

de donde se sigue inmediatamente que el estimador máximo a-posteriori para X es

$$x_{\text{MAP}} = \arg \max_{x \in \mathbb{R}^n} \{\exp(a_2(\tau)\phi(y - Tx) + a_1(\lambda)\psi(x))\} = \arg \min_{x \in \mathbb{R}^n} \{\phi(y - Tx) + \eta\psi(x)\},$$

donde $\eta \doteq a_1(\lambda)/a_2(\tau)$. Así, x_{MAP} , si existe, coincide con un minimizante de un funcional TPGG con fidelidad $\phi(y - Tx)$, penalizante $\psi(x)$ y parámetro de regularización η .

El siguiente cuadro resume los diferentes casos que hemos considerado hasta aquí, bajo un modelo lineal con ruido aditivo:

Caso	E	X	Fidelidad	Penalizante	Funcional de tipo TP
I	$\mathcal{N}(0, \tau^2 I_m)$	$\mathcal{N}(0, \lambda^2 I_n)$	$\ y - Tx\ _2^2$	$\ x\ _2^2$	$\ y - Tx\ _2^2 + \eta \ x\ _2^2$
II	$\mathcal{N}(0, \tau^2 I_m)$	$\mathcal{N}_S(\lambda, L)$	$\ y - Tx\ _2^2$	$\ Lx\ _2^2$	$\ y - Tx\ _2^2 + \eta \ Lx\ _2^2$
III	$\mathcal{N}(0, \tau^2 I_m)$	$\ell^1(\lambda)$	$\ y - Tx\ _2^2$	$\ x\ _1$	$\ y - Tx\ _2^2 + \eta \ x\ _1$
IV	$\mathcal{N}(0, \tau^2 I_m)$	$\propto \exp(a_1(\lambda) \psi(x))$	$\ y - Tx\ _2^2$	$\psi(x)$	$\ y - Tx\ _2^2 + \eta \psi(x)$
V	$\ell^1(\tau)$	$\ell^1(\lambda)$	$\ y - Tx\ _1$	$\ x\ _1$	$\ y - Tx\ _1 + \eta \ x\ _1$
VI	$\propto \exp(a_2(\tau) \phi(e))$	$\propto \exp(a_1(\lambda) \psi(x))$	$\phi(y - Tx)$	$\psi(x)$	$\phi(y - Tx) + \eta \psi(x)$

Es oportuno mencionar aquí que la existencia de un minimizante para los funcionales de tipo Tikhonov-Phillips obtenidos en los casos I, II, III y V se encuentra garantizada puesto que el término de fidelidad y el penalizante vienen dados por normas, las cuales verifican las hipótesis (G1), (H2), (H3) y (G4) del Teorema 1.27 presentado en el Capítulo 1. Para el caso IV, nuevamente, como el término de fidelidad corresponde a una norma, es fácil ver que las hipótesis (G1) y (G4) se verifican. Si además, contamos con las hipótesis (G1), (H2) y (H3) para el funcional ψ entonces, por el Teorema 1.27 se tiene la existencia de un minimizante para el funcional de Tikhonov-Phillips obtenido en dicho caso.

CASO VII.

Finalmente, consideraremos un proceso de Poisson para la variable observable $Y \in \mathbb{N}_0^m$ como en el Ejemplo 2.7. Luego, la función de verosimilitud es

$$\pi(y|x) = \prod_{i=1}^m \frac{\exp(-f_i(x)) f_i(x)^{y_i}}{y_i!}.$$

Supongamos además que X tiene una densidad a-priori como en el CASO IV dada por $\pi_{\text{pr}}(x; \lambda) = \exp(a(\lambda) \psi(x) + c(\lambda))$, donde el parámetro λ es conocido y $a(\lambda) < 0$ para todo λ . Luego, por el Teorema 2.1 se sigue inmediatamente que la densidad a-posteriori de X dado $Y = y$ es de la forma

$$\pi(x|y) \propto \left(\prod_{i=1}^m \frac{\exp(-f_i(x)) f_i(x)^{y_i}}{y_i!} \right) \exp(a(\lambda) \psi(x)).$$

Entonces, el estimador máximo a-posteriori de X está dado por

$$\begin{aligned} x_{\text{MAP}} &= \arg \max_{x \in \mathbb{R}^n} \left\{ \prod_{i=1}^m \frac{\exp(-f_i(x)) f_i(x)^{y_i}}{y_i!} \exp(a(\lambda) \psi(x)) \right\} \\ &= \arg \min_{x \in \mathbb{R}^n} \left\{ \sum_{i=1}^m (f_i(x) - y_i \ln f_i(x) + \ln y_i!) - a(\lambda) \psi(x) \right\} \end{aligned}$$

$$\begin{aligned}
&= \arg \min_{x \in \mathbb{R}^n} \left\{ \sum_{i=1}^m (f_i(x) - y_i \ln f_i(x)) - a(\lambda) \psi(x) \right\} \\
&= \arg \min_{x \in \mathbb{R}^n} \left\{ \sum_{i=1}^m (f_i(x) - y_i \ln f_i(x) + y_i \ln y_i - y_i) - a(\lambda) \psi(x) \right\} \quad (\text{si } y_i > 0 \forall i) \\
&= \arg \min_{x \in \mathbb{R}^n} \left\{ \sum_{i=1}^m \left(f_i(x) + y_i \ln \left(\frac{y_i}{f_i(x)} \right) - y_i \right) - a(\lambda) \psi(x) \right\}.
\end{aligned}$$

Por lo tanto, x_{MAP} coincide con un minimizante de un funcional TPGG con término de fidelidad dado por

$$\sum_{i=1}^m \left(f_i(x) + y_i \ln \left(\frac{y_i}{f_i(x)} \right) - y_i \right), \quad (2.7)$$

penalizante $\psi(x)$ y parámetro de regularización $\eta \doteq -a(\lambda)$. Es muy importante mencionar aquí que el término de fidelidad (2.7) corresponde a una discretización del funcional de Kullback-Leibler (1.14) presentado en el Capítulo 1 ([39]).

En conclusión, en todos los casos estudiados vimos que la función de verosimilitud y la densidad a-priori de la variable incógnita se relacionan con el término de fidelidad y el penalizante de un funcional de tipo Tikhonov-Phillips generalizado o doblemente generalizado. Más aún, vimos que el estimador máximo a-posteriori de la variable de interés coincide con un minimizante de este tipo de funcionales. Esto constituye el principal vínculo entre los enfoques determinístico y estocástico.

2.3. Una regla de elección del parámetro heurística

Es sabido que, en el método de regularización de Tikhonov-Phillips, el parámetro de regularización está asociado al compromiso entre el término de fidelidad y el de penalización, y que su elección es crucial para el éxito del método de regularización. En el Capítulo 1, hemos visto que una regla de elección del parámetro puede ser *a-priori* (depende sólo del nivel de ruido), *a-posteriori* (depende tanto del nivel de ruido como del dato) o *heurística* (sólo depende del dato con ruido). Además, hemos visto que ninguna regla de elección del parámetro heurística puede formar parte de un método de regularización convergente, aunque esto no implica que para un nivel de ruido fijo, una regla de elección del parámetro heurística no pueda funcionar de manera *adecuada*. En esta sección, estudiaremos una regla de este tipo que fue presentada por Ito et al. en [27] y cuya construcción se basa en el enfoque estocástico de resolución de

un problema inverso, más precisamente, en los modelos jerárquicos. Es por ello que, en primer lugar, veremos brevemente en qué consisten dichos modelos y luego, cómo se construye la regla en base a ellos y las características o propiedades que posee la misma.

En la sección anterior, hemos estudiado el principal vínculo existente entre los enfoques determinístico y estocástico. En particular, para el CASO I, vimos que un minimizante del funcional TP para el problema $Tx = y$ con parámetro de regularización η coincide con el estimador x_{MAP} de la distribución a-posteriori de X , cuando la densidad a-priori para esta variable está dada por $\mathcal{N}(0, \lambda^2 I_n)$, el ruido aditivo E tiene distribución $\mathcal{N}(0, \tau^2 I_m)$ y los parámetros η , λ y τ están relacionados mediante la igualdad $\eta = \frac{\tau^2}{\lambda^2}$. De esta manera, el parámetro η está determinado tanto por la varianza del error como por el parámetro λ asociado a la densidad a-priori de X . Pero, si alguno de estos parámetros son desconocidos entonces, desde el enfoque estocástico, su determinación forma parte del problema de inferencia y se debe tener en cuenta su propia distribución para la resolución, la cual a su vez puede contener parámetros desconocidos. En este caso, se obtiene lo que se denomina *modelo jerárquico* o *hipermodelo*, los cuales deben su nombre a que se establece un orden de jerarquía entre las incógnitas para resolver el problema ([9], [29]).

Consideremos ahora el modelo con ruido aditivo (2.5) y supongamos que la densidad de probabilidad del ruido E depende de uno o más parámetros, en principio desconocidos, denotados con Θ . De esta manera, la distribución del ruido está dada por $\pi_E(e|\theta)$ y la función de verosimilitud resulta

$$\pi(y|x, \theta) = \pi_E(y - f(x)|\theta).$$

Supongamos además que el vector de parámetros Θ es independiente de X y sigue una distribución $\pi_{\text{hyper}}(\theta)$. Luego, del Teorema 2.1 tenemos que la densidad a-posteriori de la dupla (X, Θ) dado $Y = y$ está dada por

$$\pi(x, \theta|y) \propto \pi(y|x, \theta) \pi(x, \theta) = \pi(y|x, \theta) \pi_{\text{pr}}(x) \pi_{\text{hyper}}(\theta).$$

De esta manera, es posible obtener la distribución a-posteriori de X y Θ a partir de la utilización de modelos jerárquicos.

Ahora, para motivar la construcción de la regla de elección del parámetro presentada en [27], consideremos el problema inverso que consiste en determinar $x \in \mathcal{X}$ en la ecuación

$$\mathcal{T}x = y, \tag{2.8}$$

donde $\mathcal{T} : \mathcal{X} \rightarrow \mathcal{Y}$ es un operador lineal y acotado entre los espacios de Banach \mathcal{X} e \mathcal{Y} con $y \in \mathcal{Y}$. Supongamos que en lugar del dato exacto y se conoce una versión con ruido $y^\delta \in \mathcal{Y}$. Como hemos mencionado previamente, en general, los problemas inversos son mal condicionados y entonces surge la necesidad de aplicar un método de regularización. Consideremos el método que consiste en la minimización de un funcional TPGG dado por

$$\mathcal{J}_{\phi, \psi, \eta}(x) = \phi(\mathcal{T}x, y) + \eta \psi(x), \quad (2.9)$$

donde $\phi : \mathcal{Y} \times \mathcal{Y} \rightarrow [0, \infty)$ y ψ es un funcional no negativo definido en \mathcal{X} . Como es bien sabido, el parámetro η está asociado al compromiso entre el término de fidelidad $\phi(\mathcal{T}x, y)$ y el penalizante $\psi(x)$. Veremos ahora como construir una regla de elección del parámetro heurística que permita estimar el valor de η . Para ello, consideremos una discretización del problema (2.8) de manera que resulte el modelo finito-dimensional

$$Tx = y^\delta, \quad (2.10)$$

con $T \in \mathbb{R}^{m \times n}$, $x \in \mathbb{R}^n$ y $y^\delta \in \mathbb{R}^m$. Abordaremos este problema desde el enfoque estadístico Bayesiano. Para ello, sea $E \in \mathbb{R}^m$ un ruido aditivo, independiente de X y con distribución dada por $\pi_E(e|\tau) = \exp\left(-\tau \widehat{\phi}(e) + c(\tau)\right)$, donde $\tau > 0$ y $\widehat{\phi} : \mathbb{R}^m \rightarrow [0, \infty)$ es una discretización del funcional $\varphi : \mathcal{Y} \rightarrow [0, \infty)$ definido como $\varphi(y_2 - y_1) \doteq \phi(y_1, y_2)$. Luego, la función de verosimilitud resulta

$$\pi(y^\delta|x, \tau) \propto \exp\left(-\tau \widehat{\phi}(y^\delta - Tx)\right).$$

Respecto de la variable de interés X , consideramos la densidad a-priori dada por

$$\pi_{\text{pr}}(x|\lambda) = \exp\left(-\lambda \widehat{\psi}(x) + C(\lambda)\right),$$

donde $\lambda > 0$ y $\widehat{\psi} : \mathbb{R}^n \rightarrow [0, \infty)$ es una discretización del funcional ψ . Es oportuno mencionar aquí que las densidades propuestas para X y E están asociadas a la familia exponencial (2.6), siendo un caso particular de aquellas densidades consideradas en el CASO VI de la Sección 2.2 para estas variables (aquí, $a_1(\lambda) = -\lambda$, $a_2(\tau) = -\tau$, $c_1(\lambda) = C(\lambda)$ y $c_2(\tau) = c(\tau)$).

Supongamos ahora que los parámetros λ y τ son desconocidos y, por ende, son considerados variables aleatorias. Si $\lambda \sim \Gamma(\alpha_0, \beta_0)$ y $\tau \sim \Gamma(\alpha_1, \beta_1)$, sus densidades son de la forma

$$\pi(\lambda) \propto \lambda^{\alpha_0-1} e^{-\beta_0 \lambda} \quad y \quad \pi(\tau) \propto \tau^{\alpha_1-1} e^{-\beta_1 \tau},$$

respectivamente. Aquí, (α_0, β_0) y (α_1, β_1) son pares de parámetros conocidos.

Como consecuencia del Teorema 2.1, la densidad a-posteriori de la terna (X, Λ, Υ) resulta

$$\begin{aligned} \pi(x, \lambda, \tau | y^\delta) &\propto \pi(y^\delta | x, \tau) \pi(x | \lambda) \pi(\lambda) \pi(\tau) \\ &\propto \exp\left(-\tau \widehat{\phi}(y^\delta - Tx) - \lambda \widehat{\psi}(x) + (\alpha_0 - 1) \ln \lambda + (\alpha_1 - 1) \ln \tau - \beta_0 \lambda - \beta_1 \tau\right), \end{aligned}$$

de donde se sigue que el estimador máximo a-posteriori para (X, Λ, Υ) está dado por aquella terna que minimiza el funcional

$$J(x, \lambda, \tau) = \tau \widehat{\phi}(y^\delta - Tx) + \lambda \widehat{\psi}(x) + \beta_0 \lambda - (\alpha_0 - 1) \ln \lambda + \beta_1 \tau - (\alpha_1 - 1) \ln \tau.$$

Finalmente, el funcional J motiva el *funcional de Tikhonov-Phillips generalizado aumentado* ([27], [28]) dado por

$$\mathcal{J}(x, \lambda, \tau) = \tau \phi(Tx, y^\delta) + \lambda \psi(x) + \beta_0 \lambda - \alpha_0 \ln \lambda + \beta_1 \tau - \alpha_1 \ln \tau, \quad (2.11)$$

definido para $(x, \lambda, \tau) \in \mathcal{X} \times \mathbb{R}^+ \times \mathbb{R}^+$, que extiende al funcional (2.9). Cabe mencionar aquí que se utilizó la misma notación para los parámetros α_0 , α_1 , β_0 y β_1 con el objetivo de simplificar la notación. A su vez, es oportuno aclarar que el funcional $\mathcal{J}(x, \lambda, \tau)$ no será utilizado para resolver el problema (2.8) sino que se introduce sólo para lograr construir una regla para la elección de un parámetro de regularización η *razonable* en (2.9). Recordemos que el principal objetivo es estimar la variable incógnita X y, para ello, consideraremos su estimador x_{MAP} que es aquel que minimiza el funcional $\mathcal{J}(x, \lambda, \tau)$ en (2.11). Con este objetivo entonces, en [27] se construye la regla de elección del parámetro a partir del concepto de *punto crítico* del funcional $\mathcal{J}(x, \lambda, \tau)$ que veremos a continuación.

Definición 2.9. Un elemento $(x^*, \lambda^*, \tau^*) \in \mathcal{X} \times \mathbb{R}^+ \times \mathbb{R}^+$ es un *punto crítico* del funcional (2.11) si satisface el siguiente sistema de optimalidad generalizado

$$\begin{cases} x^* = \arg \min_{x \in \mathcal{X}} \{ \phi(Tx, y^\delta) + \lambda^*(\tau^*)^{-1} \psi(x) \} \\ 0 = \psi(x^*) + \beta_0 - \alpha_0 \frac{1}{\lambda^*} \\ 0 = \phi(Tx^*, y^\delta) + \beta_1 - \alpha_1 \frac{1}{\tau^*}. \end{cases}$$

Es importante observar que esta definición de punto crítico coincide con la que se realiza habitualmente anulando las derivadas parciales $\frac{\partial}{\partial \lambda} \mathcal{J}(x^*, \lambda^*, \tau^*) = \psi(x^*) + \beta_0 - \alpha_0 \frac{1}{\lambda^*}$ y $\frac{\partial}{\partial \tau} \mathcal{J}(x^*, \lambda^*, \tau^*) = \phi(Tx^*, y^\delta) + \beta_1 - \alpha_1 \frac{1}{\tau^*}$. De allí, resultan

$$\lambda^* = \frac{\alpha_0}{\psi(x^*) + \beta_0} \quad \text{y} \quad \tau^* = \frac{\alpha_1}{\phi(Tx^*, y^\delta) + \beta_1}.$$

Por lo tanto, de la Definición 2.9 se sigue que el parámetro de regularización η^* está determinado por

$$\eta^* \doteq \lambda^* (\tau^*)^{-1} = \frac{\alpha_0}{\psi(x_{\eta^*}^\delta) + \beta_0} \frac{\phi(\mathcal{T}x_{\eta^*}^\delta, y^\delta) + \beta_1}{\alpha_1}, \quad (2.12)$$

donde denotamos con $x_{\eta^*}^\delta$ a x^* .

Observación 2.10. Es oportuno mencionar aquí que es posible construir una regla más general que la de la ecuación (2.12) a partir del concepto de punto crítico del funcional

$$\mathcal{J}(x, \lambda, \tau) = -a_2(\tau) \phi(\mathcal{T}x, y^\delta) - a_1(\lambda) \psi(x) + \beta_0 \lambda - \alpha_0 \ln \lambda + \beta_1 \tau - \alpha_1 \ln \tau,$$

que se obtiene considerando que las variables X y E tienen distribuciones como las dadas en el CASO VI de la Sección 2.2. Esto constituye un problema abierto y un posible trabajo futuro.

De la teoría determinística de regularización, se sabe que una regla de elección del parámetro a-priori $\eta(\delta)$ debe verificar que

$$\lim_{\delta \rightarrow 0^+} \eta(\delta) = 0 \quad y \quad \lim_{\delta \rightarrow 0^+} \frac{\delta^2}{\eta(\delta)} = 0, \quad (2.13)$$

para asegurar que la solución aproximada converja a una solución de mínimo penalizante cuando el nivel de ruido tiende a cero (Proposición 1.11, Capítulo 1). Por otro lado, sabemos que en el contexto estocástico el nivel de ruido está asociado a la variable E . Más aún, para el caso en que $E \sim \mathcal{N}(0, \sigma^2 I)$, el nivel de ruido δ se encuentra estrechamente asociado con la varianza σ^2 del error. De esta forma, pensar en “ $\delta \rightarrow 0^+$ ” en el contexto determinístico es análogo a pensar en “ $\sigma \rightarrow 0^+$ ”, en el estocástico. Además, es sencillo ver que para este último caso existe una relación entre σ y τ dada por $\sigma^2 = \frac{1}{\tau}$. Debido a que el segundo límite en (2.13) no se verifica para la regla en (2.12), en [27] se propone modificar esta regla de la siguiente manera

$$\eta^* = \frac{\alpha_0}{\psi(x_{\eta^*}^\delta) + \beta_0} \left(\frac{\phi(\mathcal{T}x_{\eta^*}^\delta, y^\delta)}{\alpha_1} \right)^{-d} \frac{\phi(\mathcal{T}x_{\eta^*}^\delta, y^\delta)}{\alpha_1} = \alpha \frac{\phi(\mathcal{T}x_{\eta^*}^\delta, y^\delta)^{1-d}}{\psi(x_{\eta^*}^\delta) + \beta_0}, \quad (2.14)$$

donde $0 < d < 1$, $\alpha = \frac{\alpha_0}{(\alpha_1)^{1-d}}$ y β_1 se descarta por considerarse pequeño. Cabe mencionar aquí que experiencias numéricas han mostrado que la elección $d = 0$ (correspondiente a la regla original (2.12)) no funciona de manera adecuada en ciertos casos.

Finalmente, la regla de elección del parámetro propuesta consiste en encontrar el parámetro η que verifica la siguiente ecuación no lineal

$$\eta \left(\psi(x_\eta^\delta) + \beta_0 \right) = \alpha \phi(\mathcal{T}x_\eta^\delta, y^\delta)^{1-d}, \quad 0 < d < 1. \quad (2.15)$$

A continuación, presentamos algunos resultados importantes sobre esta regla, cuyas pruebas pueden consultarse en [27]. En primer lugar, enunciaremos un resultado que asegura la existencia de una solución para la ecuación (2.15). Luego, presentaremos un algoritmo numérico para su obtención y un resultado que asegura su convergencia para el cual es necesario definir previamente cierta propiedad que puede satisfacer el penalizante. Para todos estos resultados, son necesarias las siguientes hipótesis:

- I. \mathcal{X} es un espacio reflexivo.
- II. Para cada $\eta > 0$, el funcional $\mathcal{J}_{\phi,\psi,\eta}(x)$ definido en (2.9) es coercivo en \mathcal{X} , es decir, $\mathcal{J}_{\phi,\psi,\eta}(x) \rightarrow +\infty$ si $\|x\| \rightarrow +\infty$.
- III. Los funcionales $\phi(\mathcal{T}x, y)$ y $\psi(x)$ son débilmente semicontinuos inferiormente respecto de x , es decir, para cada sucesión $\{x_n\}_n \subset \mathcal{X}$ tal que $x_n \rightharpoonup x$ se tiene que $\phi(\mathcal{T}x, y) \leq \liminf_{n \rightarrow \infty} \phi(\mathcal{T}x_n, y)$ y $\psi(x) \leq \liminf_{n \rightarrow \infty} \psi(x_n)$.
- IV. Existe $\tilde{x} \in \mathcal{X}$ tal que $\psi(\tilde{x}) = 0$.

Cabe destacar aquí que las Hipótesis II y III son similares a las usuales para asegurar la existencia de un minimizante para el funcional de Tikhonov-Phillips $\mathcal{J}_{\phi,\psi,\eta}$, como hemos visto en el Capítulo 1.

El siguiente resultado garantiza la existencia de una solución para la ecuación no lineal (2.15), es decir, garantiza la existencia de un parámetro óptimo η^* .

Teorema 2.11. Sean $\phi(\mathcal{T}x_\eta^\delta, y)$ y $\psi(x_\eta^\delta)$ funcionales continuos respecto de η y supongamos que $\lim_{\eta \rightarrow 0^+} \phi(\mathcal{T}x_\eta^\delta, y) > 0$ entonces existe al menos una solución positiva de la ecuación (2.15).

Para implementar numéricamente la regla de elección de parámetro heurística (2.15), se propone el siguiente algoritmo que, de manera iterada, obtiene el parámetro de regularización y la solución regularizada correspondiente a dicho parámetro.

Algoritmo: Elegir una posición inicial $\eta_0 > 0$ y fijar $k = 0$. Hallar (x_k, η_k) para $k \geq 1$ de la siguiente manera:

1. Hallar el minimizante del funcional TPGG dado por

$$x_{k+1} = \arg \min_x \left\{ \phi(\mathcal{T}x, y^\delta) + \eta_k \psi(x) \right\}.$$

2. Actualizar el parámetro η_{k+1} como $\eta_{k+1} = \alpha \frac{\phi(\mathcal{T}x_{k+1}, y^\delta)^{1-d}}{\psi(x_{k+1}) + \beta_0}$.
3. Chequear el criterio de parada establecido. Si no se alcanza, fijar $k = k + 1$ y repetir los pasos 1 - 3.

Como criterio de parada puede establecerse, por ejemplo, que el algoritmo finalice cuando la diferencia entre dos iteraciones sucesivas no supere cierto umbral, ya sea para las iteraciones en x como en η .

Definición 2.12. Un funcional ψ se dice que *tiene la H-propiedad* en el espacio \mathcal{X} si para cada sucesión $\{x_n\}_n \subset \mathcal{X}$ que converge débilmente a $x_0 \in \mathcal{X}$ y tal que $\psi(x_n) \rightarrow \psi(x_0)$, se verifica que la sucesión $\{x_n\}_n$ también converge fuertemente a x_0 en \mathcal{X} .

Teorema 2.13. *Sea $\eta^* > 0$. Entonces, cada subsucesión de $\{(x_k, \eta_k)\}_k$ generada por el algoritmo tiene una subsucesión débilmente convergente a una solución (x^*, η^*) de la ecuación (2.15), y la convergencia de la sucesión $\{\eta_k\}_k$ es monótona. Si el minimizante del funcional $\mathcal{J}_{\phi, \psi, \eta^*}(x)$ es único, entonces toda la sucesión converge débilmente. Además, si el funcional ψ tiene la H-propiedad, entonces la convergencia es fuerte.*

Es oportuno mencionar aquí que la H-propiedad que satisface el funcional ψ en el teorema anterior es análoga a aquella presentada en la Observación 1.34 para garantizar la convergencia fuerte de soluciones regularizadas obtenidas a partir de la minimización de funcionales de tipo TPGG a una solución ψ -minimizante, cuando el nivel de ruido tendía a cero por derecha.

En la siguiente sección, presentaremos diferentes ejemplos numéricos en donde se implementó el algoritmo presentado para obtener un parámetro de regularización óptimo en problemas inversos de procesamiento de señales e imágenes.

2.4. Aplicaciones

Si bien no es un objetivo principal de este trabajo de tesis el abordaje numérico de los problemas inversos, consideramos importante presentar algunos ejemplos numéricos en donde se implementó la regla de elección del parámetro estudiada en la Sección 2.3 para obtener un parámetro de regularización óptimo en problemas inversos de procesamiento de señales e imágenes. En todos los ejemplos consideramos el término de fidelidad cuadrático $\phi(\mathcal{T}x, y) = \frac{1}{2} \|\mathcal{T}x - y\|^2$. Además, para el caso de procesamiento de imágenes, veremos un ejemplo de modelo jerárquico que nos permitirá detectar fuentes focales presentes en la imagen original sin la utilización de una regla de elección del parámetro.

2.4.1. Señales

Un modelo matemático básico para la convolución de señales está dado por la ecuación integral de Fredholm de primera clase

$$y(s) = \int_I a(s, t) x(t) dt,$$

donde a corresponde al núcleo de convolución, x es la señal que se desea aproximar, y es el dato del problema y la integración se realiza sobre cierto intervalo real I . Luego de discretizar la ecuación integral anterior de manera usual (usando colocación y cuadratura), se obtiene un modelo de la forma $y = Tx$, donde $x, y \in \mathbb{R}^n$ y la matriz $T \in \mathbb{R}^{n \times n}$ resulta de la discretización del núcleo a . En los ejemplos presentados a continuación, se consideró $n = 100$ y el dato fue contaminado mediante un ruido aditivo Gaussiano de la siguiente manera

$$y_i^\delta = y_i + \varepsilon \max_i |y_i| \xi_i,$$

donde ε es un parámetro positivo conocido y $\xi_i \sim \mathcal{N}(0, 1)$. Como consecuencia, existe una relación entre el parámetro ε y la varianza del error σ^2 dada por $\sigma = \varepsilon \max_i |y_i|$.

Ejemplo 2.14. Este ejemplo se basa en uno de los presentados en [27], en el cual se considera el núcleo de convolución a dado por $a(s, t) \doteq \varphi(s - t)$, con $\varphi(t) = (1 + \cos(\frac{\pi}{3}t)) \chi_{|t-s| < 3}$, donde χ denota la función característica. Consideremos además que la señal original viene dada por

$$x(t) \doteq -1,5 \chi_{[-5, 12; 5]} + \chi_{[-0, 92; -0, 8]} - 1,2 \chi_{[0, 9; 1, 02]} - \chi_{[4, 1; 4, 22]},$$

y el intervalo de integración correspondiente es $I = [-6, 6]$. Debido a que la señal original consta de *saltos o discontinuidades*, se opta por utilizar el término penalizante $\psi(x) \doteq \|x\|_1$. Como hemos visto en la Sección 2.1.1 dicho penalizante se encuentra estrechamente relacionado con la densidad a-priori de impulso ℓ^1 , que es utilizada para modelar este tipo de comportamiento.

En la Figura 5, se muestra el dato contaminado con ruido aditivo Gaussiano mediante el procedimiento descrito previamente, con $\varepsilon = 0,01$.

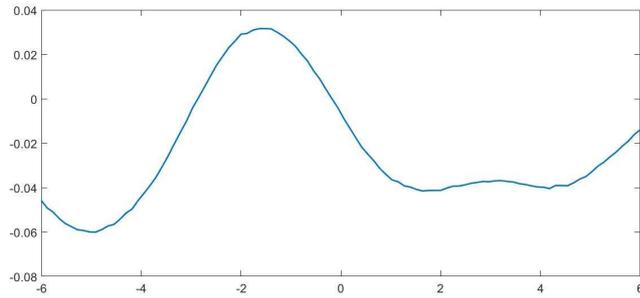


FIGURA 5. Dato contaminado con ruido aditivo Gaussiano, $\varepsilon = 0,01$.

Mediante la regla de elección del parámetro presentada en la Sección 2.3, utilizando los valores $\eta_0 = 10^{-5}$; $\alpha = 0,1$; $d = 1/3$ y $\beta_0 = 10^{-4}$ y luego de 6 iteraciones, se obtuvo el parámetro de regularización óptimo $\eta^* = 0,82 \times 10^{-5}$. En la Figura 6, se presentan la señal original x y la solución $x_{\eta^*}^\delta$ obtenida. Es posible concluir que el procedimiento implementado logra recuperar las principales características de la señal original.

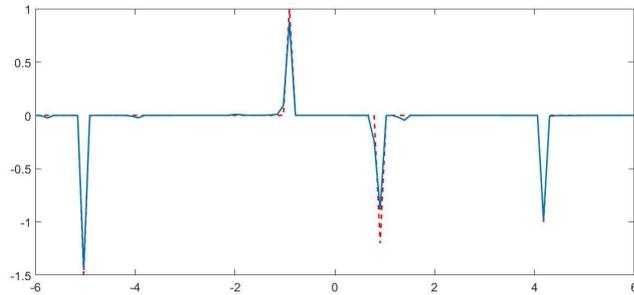


FIGURA 6. Señal original (---) y solución regularizada (—).

Para medir el error en la estimación de X utilizaremos el error relativo dado por $\frac{\|x_{\eta^*}^\delta - x\|_2}{\|x\|_2}$ que indica la proporción del error absoluto con respecto al valor exacto de $\|x\|_2$. El error relativo e obtenido para este caso fue $e = 0,1534$.

La siguiente tabla presenta el promedio de 30 errores relativos obtenidos de repetir el proceso de obtención del parámetro óptimo de regularización para cada valor de ε .

ε	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
e	0,201	0,054	0,008	0,0022	$7,10 \cdot 10^{-4}$

Ejemplo 2.15. Este ejemplo se basa en uno de los que se encuentran en [27], donde se utiliza el núcleo de convolución $a(s, t) \doteq (\cos s + \cos t) \left(\frac{\sin u}{u}\right)^2$, con $u(s, t) = \pi(\sin s + \sin t)$ y la

señal original está dada por $x(t) \doteq 2e^{-6(t-4/5)^2} + e^{-2(t+1/2)^2}$, definida en el intervalo $I = [-\pi/2, \pi/2]$. Puesto que la señal original es suave, se opta por considerar el término penalizante $\psi(x) \doteq \frac{1}{2} \|x\|^2$, el cual, como vimos en la Sección 2.1.1, se corresponde con una densidad a-priori Gaussiana.

En la Figura 7, se muestra el dato contaminado con ruido aditivo Gaussiano a partir del procedimiento descrito previamente, utilizando $\varepsilon = 0,01$.

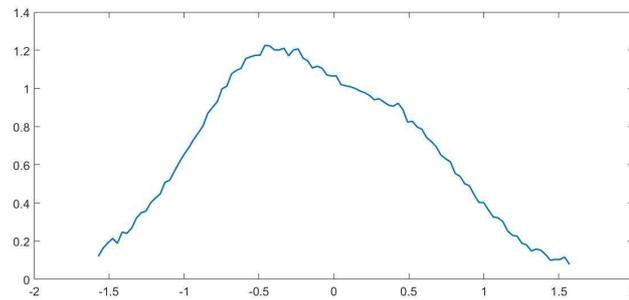


FIGURA 7. Dato contaminado con ruido aditivo Gaussiano, $\varepsilon = 0,01$.

A partir de la regla de elección del parámetro presentada en la Sección 2.3, con $\eta_0 = 10^{-5}$, $\alpha = 1$, $d = 1/3$ y $\beta_0 = 10^{-4}$ y luego de 9 iteraciones, se obtuvo el parámetro de regularización óptimo $\eta^* = 0,724 \times 10^{-3}$. En la Figura 8, se presentan la señal original x y la solución $x_{\eta^*}^\delta$ obtenida, la cual constituye una buena aproximación de la señal original.

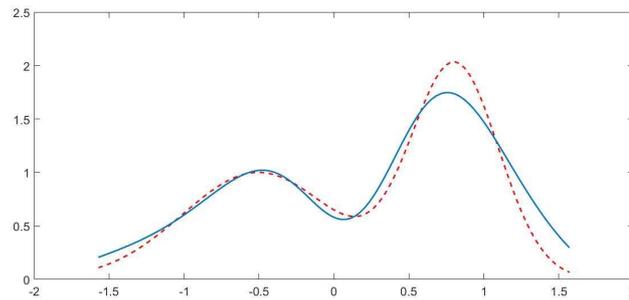


FIGURA 8. Señal original (---) y solución regularizada (—).

El error relativo e dado por $\frac{\|x_{\eta^*}^\delta - x\|_2}{\|x\|_2}$ obtenido para este caso fue $e = 0,147$.

La siguiente tabla presenta, análogamente al ejemplo anterior, el promedio de 30 errores relativos obtenidos de repetir el proceso de obtención del parámetro de regularización óptimo para cada valor de ε .

ε	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
e	0,141	0,058	0,045	0,033	0,028

En ambos ejemplos se observa, como es de esperar, que a medida que se perturba menos la señal original, los errores relativos disminuyen.

Ejemplo 2.16. Para la señal original x del Ejemplo 2.14, si repetimos el proceso pero ahora considerando el penalizante $\psi(x) = \frac{1}{2}\|x\|^2$, se obtiene la solución regularizada presente en la Figura 9. Debido a que el penalizante utilizado fuerza una solución regularizada suave, la misma intenta reproducir el comportamiento de la señal original pero respetando la “suavidad” que exige el penalizante.

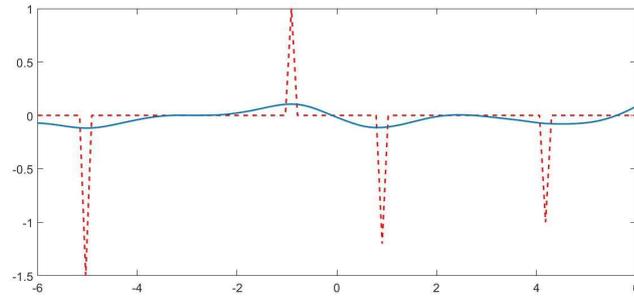


FIGURA 9. Señal original (---) y solución regularizada (—) para $\psi(x) = \frac{1}{2}\|x\|^2$.

Análogamente, si en el Ejemplo 2.15 consideramos el penalizante $\psi(x) = \|x\|_1$ y se repite el proceso, se obtiene la solución regularizada que se presenta en la Figura 10. Cabe mencionar aquí que el método fuerza una solución regularizada con “picos” debido al uso de un penalizante que es adecuado para detectar discontinuidades o falta de regularidad de la solución.

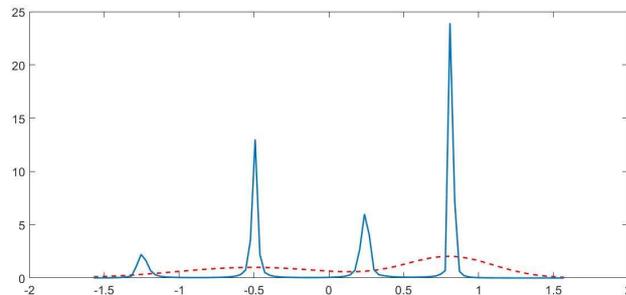


FIGURA 10. Señal original (---) y solución regularizada (—) para $\psi(x) = \|x\|_1$.

En conclusión, si bien el parámetro de regularización es elegido de manera óptima mediante la regla (2.14), la elección de un penalizante apropiado resulta crucial para el éxito del método.

2.4.2. Imágenes

Un modelo matemático muy general para el proceso de degradación de imágenes está dado por la siguiente ecuación integral

$$g(s, t) = \int_0^1 \int_0^1 a(s, t, s', t') f(s', t') ds' dt', \quad (2.16)$$

donde f y g son las imágenes original y degradada, respectivamente, y a corresponde al núcleo de dispersión. Para los ejemplos numéricos presentados en esta sección utilizaremos el núcleo de convolución Gaussiano dado por

$$a(s, t, s', t') = \frac{1}{2\pi\sigma_h\sigma_v} \exp\left(-\frac{1}{2}\left(\frac{s-s'}{\sigma_h}\right)^2 - \frac{1}{2}\left(\frac{t-t'}{\sigma_v}\right)^2\right), \quad (2.17)$$

donde las desviaciones estándar σ_h y σ_v caracterizan la degradación en la dirección horizontal y vertical, respectivamente. La discretización del modelo (2.16) con núcleo Gaussiano y con ruido aditivo \tilde{e} resulta en $\tilde{y} = T\tilde{x} + \tilde{e}$, donde la matriz T , de orden $N \times N$, está asociada a la discretización del núcleo a y $N = n^2$, siendo la imagen original de tamaño $n \times n$. Luego, la versión estocástica del modelo discretizado está dada por

$$Y = TX + E, \quad (2.18)$$

donde $Y, X, E \in \mathbb{R}^N$ son vectores aleatorios. El objetivo es, entonces, estimar X utilizando su distribución a-posteriori construida a partir del dato $Y = y$ y de la información a-priori disponible de la variable incógnita.

Supongamos que la imagen original consta de un fondo negro con algunos objetos focales, como por ejemplo estrellas, cuya ubicación y forma queremos determinar. Si bien el dato puede aportar una sugerencia respecto de la localización, es posible que una región borrosa en el dato contenga más de una fuente focal en la imagen original. Para modelar este tipo de variable, en primer lugar, consideremos una densidad a-priori de impulso $\ell^1(\alpha)$ con restricción de positividad cuya densidad está dada por (2.4), donde el parámetro α se supone conocido. Supongamos además que el ruido $E \sim \mathcal{N}(0, \sigma_e^2 I_N)$ (σ_e conocido) e independiente de X de donde resulta que la función de verosimilitud es de la forma

$$\pi(y|x) \propto \exp\left(-\frac{1}{2\sigma_e^2}\|y - Tx\|_2^2\right).$$

Finalmente, por el Teorema 2.1 se sigue que

$$\pi(x|y) \propto \pi_+(x) \exp\left(-\frac{1}{2\sigma_e^2}\|y - Tx\|_2^2 - \alpha\|x\|_1\right). \quad (2.19)$$

Es oportuno recordar que nuestro principal objetivo es estimar la variable incógnita X y lo haremos a partir del estimador máximo a-posteriori. Es decir,

$$x_{\text{MAP}} = \arg \min_x \left\{ \frac{1}{\sigma_e^2} \|y - Tx\|_2^2 + \alpha \|x\|_1 \right\}, \quad (2.20)$$

de donde se sigue inmediatamente que x_{MAP} es un minimizante de un funcional TPG con penalizante $\|x\|_1$ y parámetro de regularización $\eta \doteq \sigma_e^2 \alpha$.

En el siguiente ejemplo, supondremos que la imagen original es de 60×60 píxeles y consta de cuatro fuentes focales que se distinguen de un fondo negro. Esta imagen es degradada mediante el núcleo Gaussiano (2.17) utilizando $\sigma_h = \sigma_v = 2$ y condición de borde reflexiva ([21]), y se considera ruido aditivo $E \sim \mathcal{N}(0, \sigma_e^2 I_N)$ con $\sigma_e = 0,01 \|y\|_\infty$ y $N = 60^2$. En la Figura 11, se muestran las imágenes original y borrosa con ruido.

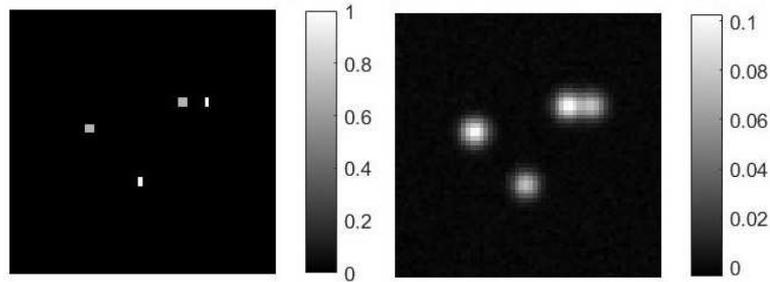


FIGURA 11. Imagen original (izquierda) y borrosa con ruido (derecha).

Mediante la regla de elección del parámetro presentada en la Sección 2.3, utilizando los parámetros $\eta_0 = 10^{-5}$, $\alpha = 0,05$; $d = 1/3$ y $\beta_0 = 10^{-10}$, y luego de 6 iteraciones, se obtuvo el parámetro de regularización óptimo $\eta^* = 1,2135 \times 10^{-4}$ y la solución que se muestra en la Figura 12.

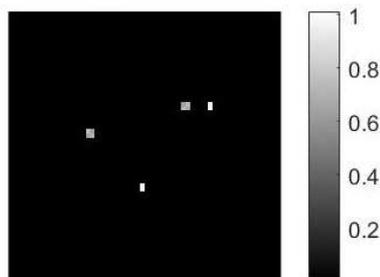


FIGURA 12. Solución obtenida con $\eta^* = 1,2135 \times 10^{-4}$.

Es posible concluir que el método implementado logra detectar la ubicación y la forma de los objetos focales presentes en la imagen original. El error relativo e dado por $\frac{\|x_{\eta^*}^\delta - x\|_2}{\|x\|_2}$ obtenido para este caso fue $e = 0,0555$.

La siguiente tabla presenta, análogamente a lo realizado en procesamiento de señales, el promedio de 30 errores relativos obtenidos de repetir el proceso de obtención del parámetro óptimo de regularización para cada valor de ε .

ε	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
e	0,055	0,015	0,001	$1,164 \cdot 10^{-4}$	$6,076 \cdot 10^{-5}$

Finalmente, mostraremos cómo se podrían utilizar modelos jerárquicos para detectar la ubicación y la forma de los objetos focales presentes en la imagen original dada en la Figura 11. Para esto, consideraremos que X tiene una densidad a-priori Gaussiana de media cero y matriz de covarianzas $D_\theta \doteq \text{diag}(\theta_1, \theta_2, \dots, \theta_N)$, es decir

$$\pi_{pr}(x) \propto \left(\frac{1}{\theta_1 \theta_2 \dots \theta_N} \right)^{1/2} \exp \left(-\frac{1}{2} \sum_{j=1}^N \frac{x_j^2}{\theta_j} \right) = \exp \left(-\frac{1}{2} \|D_\theta^{-1/2} x\|_2^2 - \frac{1}{2} \sum_{j=1}^N \log \theta_j \right).$$

Intuitivamente, la varianza θ_j podría ser grande para un píxel correspondiente a un objeto focal, mientras que debería ser pequeña para los píxeles correspondientes al fondo negro. De esta manera, esperamos que la cantidad de píxeles no nulos sea pequeña o que dichos píxeles puedan ser interpretados como datos atípicos. Es por esta razón que supondremos que el vector θ asociado a la matriz de covarianzas de X tiene una distribución Gamma Inversa de parámetros (α, θ_0) conocidos, dando lugar a un modelo jerárquico. Análogamente al ejemplo anterior, supondremos $E \sim \mathcal{N}(0, \sigma_e^2 I_N)$ (σ_e conocido). Con las densidades de los vectores involucrados, obtenemos una densidad conjunta de la forma

$$\pi(x, \theta, y) \propto \exp \left(-\frac{1}{2\sigma_e^2} \|Tx - y\|_2^2 - \frac{1}{2} \|D_\theta^{-1/2} x\|_2^2 - \sum_{j=1}^N \frac{\theta_0}{\theta_j} - \left(\alpha + \frac{3}{2} \right) \sum_{j=1}^N \log \theta_j \right).$$

Luego,

$$\begin{aligned} \pi(x|\theta, y) &\propto \exp \left(-\frac{1}{2\sigma_e^2} \|Tx - y\|_2^2 - \frac{1}{2} \|D_\theta^{-1/2} x\|_2^2 \right) \\ \pi(\theta|x, y) &\propto \exp \left(-\frac{1}{2} \|D_\theta^{-1/2} x\|_2^2 - \sum_{j=1}^N \frac{\theta_0}{\theta_j} - \left(\alpha + \frac{3}{2} \right) \sum_{j=1}^N \log \theta_j \right). \end{aligned}$$

Para obtener los estimadores x_{MAP} y θ_{MAP} , se utiliza un algoritmo iterativo de manera tal que dada la estimación $(x^{(k)}, \theta^{(k)})$ en el paso k , se obtiene:

Paso 1: $x^{(k+1)} = \arg \max \pi(x|\theta^{(k)}, y)$, resolviendo el sistema $\begin{pmatrix} T \\ \sigma_e D_{\theta^{(k)}}^{-1/2} \end{pmatrix} x = \begin{pmatrix} y \\ 0 \end{pmatrix}$;

Paso 2: $\theta^{(k+1)} = \arg \max \pi(\theta|x^{(k+1)}, y)$, dado por $\theta_j^{(k+1)} = \frac{1}{\alpha + 3/2} \left(\frac{1}{2} (x_j^2)^{(k+1)} + \theta_0 \right)$.

Utilizando los parámetros $\alpha = 1000$, $\theta_0 = 1$, luego de 300 iteraciones obtenemos el estimador θ_{MAP} para θ que se muestra en la Figura 13.

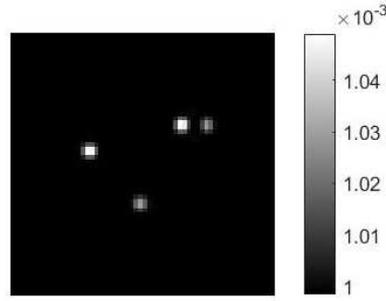


FIGURA 13. Estimador θ_{MAP} para θ .

En la Figura 14 se presenta el estimador x_{MAP} obtenido para θ_{MAP} en donde puede observarse que no se logra recuperar ni la forma ni la intensidad de las fuentes focales presentes en la imagen original.

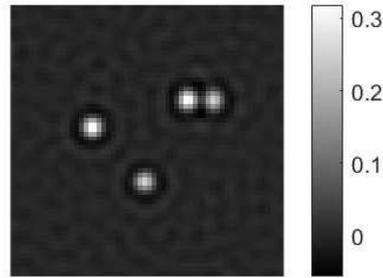


FIGURA 14. Estimador x_{MAP} para X a partir de θ_{MAP} .

Debido a que las componentes del vector θ están asociadas a las varianzas de los píxeles en X , optamos por construir una imagen binaria $\tilde{\theta}$ a partir de θ_{MAP} que tenga en cuenta sólo aquellos píxeles con intensidades mayores que cierto umbral M . Así, se construye $\tilde{\theta}$ como $\tilde{\theta}(i, j) = \begin{cases} 1, & \theta_{\text{MAP}}(i, j) \geq M \\ k, & \theta_{\text{MAP}}(i, j) < M \end{cases}$, donde $M = \frac{\max \theta_{\text{MAP}} + \min \theta_{\text{MAP}}}{2}$ y k es cercano a cero, obteniendo la imagen que se muestra en la Figura 15.

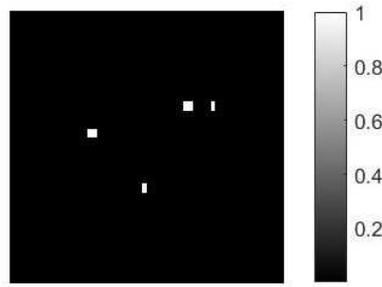


FIGURA 15. Imagen binaria $\tilde{\theta}$ construída a partir del estimador θ_{MAP} .

Finalmente, se realizó el Paso 1 del algoritmo utilizando $\tilde{\theta}$ y se obtuvo el estimador x_{MAP} para X que se presenta en la Figura 16.

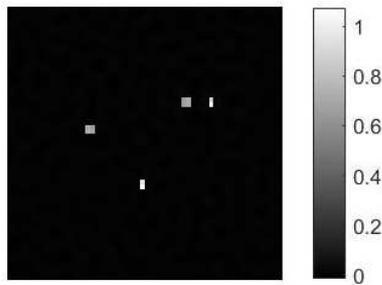


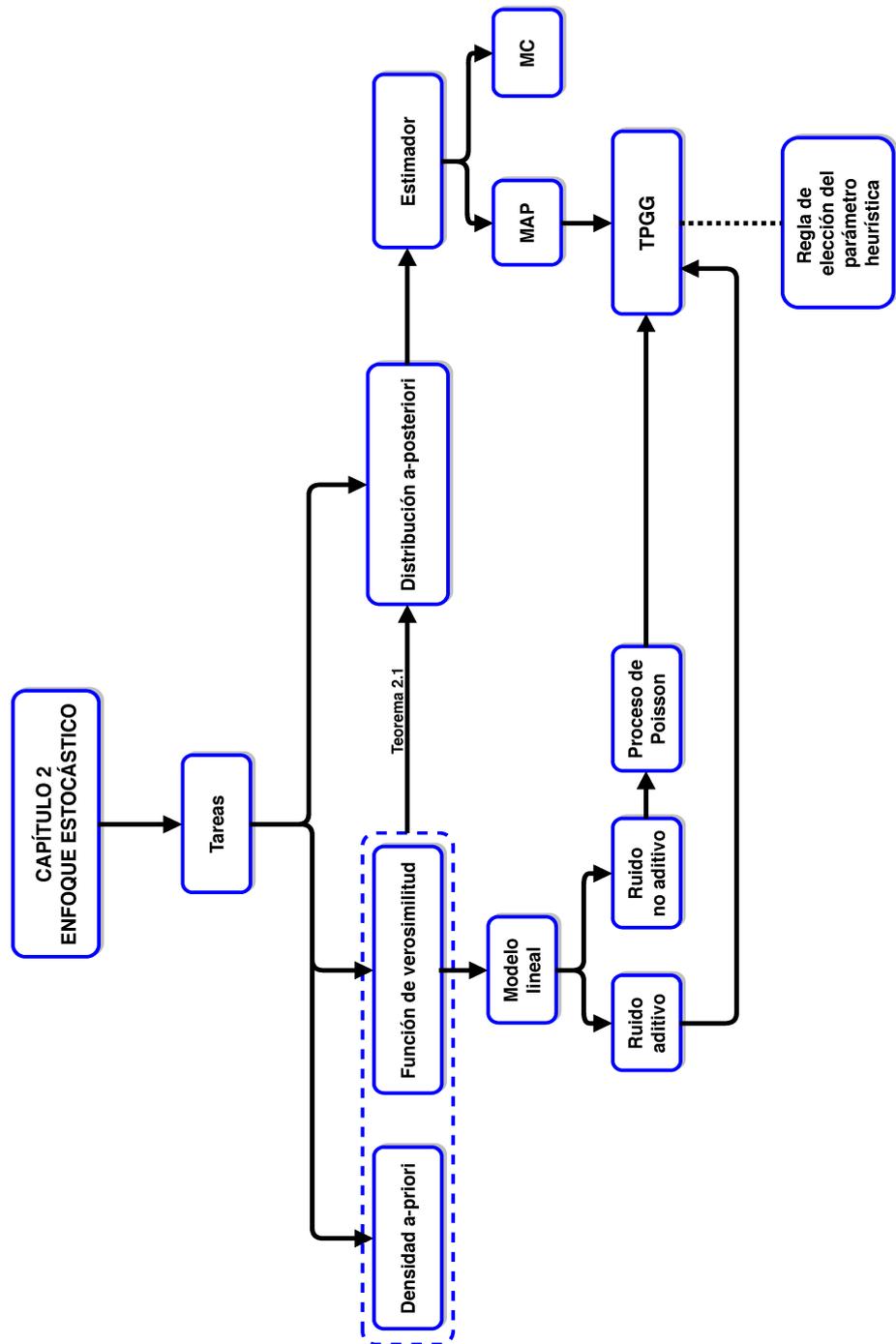
FIGURA 16. Estimador x_{MAP} para X a partir de $\tilde{\theta}$.

El error relativo e dado por $\frac{\|x_{\text{MAP}} - x\|_2}{\|x\|_2}$ obtenido para este caso fue $e = 0,1117$. Cabe mencionar aquí que, a pesar de que luego del umbralado las intensidades de los píxeles correspondientes a las fuentes focales en $\tilde{\theta}$ son las mismas (e iguales a 1), el estimador x_{MAP} logra recuperar las intensidades originales de dichos píxeles.

La siguiente tabla presenta el promedio de 30 errores relativos obtenidos para cada valor de ε tras repetir el procedimiento descrito.

ε	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
e	0,111	0,041	0,0289	0,0288	0,0286

En conclusión, a lo largo de esta sección hemos presentado algunos ejemplos numéricos en aplicaciones de procesamiento de señales e imágenes mostrando el comportamiento de la regla heurística 2.15. Asimismo, hemos puesto de manifiesto la importancia de una elección adecuada del término penalizante para el éxito del método. Finalmente, hemos aplicado un modelo jerárquico para el ejemplo presentado de restauración de imágenes que logra recuperar las fuentes focales presentes en la imagen original.



CAPÍTULO 3

Estimaciones del error a-posteriori

Consideremos nuevamente el problema inverso que consiste en determinar x en la ecuación

$$\mathcal{T}x = y, \tag{3.1}$$

donde $\mathcal{T} : \mathcal{X} \rightarrow \mathcal{Y}$ es un operador lineal y acotado entre espacios de Banach de dimensión infinita tal que $\mathcal{R}(\mathcal{T})$ es no cerrado. Siguiendo el enfoque determinístico para su resolución, en el Capítulo 1 vimos que, cuando un problema inverso resulta mal condicionado en el sentido de Hadamard debido a la falta de cumplimiento del postulado (Had 3), es necesario aplicar métodos de regularización. En particular, estudiamos el método de regularización de Tikhonov-Phillips en su versión clásica y propusimos dos versiones generalizadas del mismo que consisten en la minimización de funcionales TPG y TPGG. Pudimos probar que estos métodos propuestos constituían regularizaciones en el contexto de espacios de Banach. Para ello, fue necesario probar la convergencia de las soluciones regularizadas a una solución de mínimo penalizante. Esto lo hicimos para el caso de reglas de elección del parámetro a-priori.

En el Capítulo 2, vimos en qué consiste resolver un problema inverso desde el enfoque estocástico y presentamos una regla de elección del parámetro heurística cuya construcción se basa en este nuevo enfoque y en su relación con el funcional de Tikhonov-Phillips “aumentado”. Como hemos mencionado, es sabido que este tipo de reglas no pueden formar parte de un método de regularización convergente pero esto no indica que la misma no pueda tener un buen desempeño para un nivel de ruido pequeño y fijo. En este sentido entonces, en el presente capítulo nos proponemos encontrar estimaciones para el error a-posteriori, es decir, encontrar una cota para el error entre una solución de mínimo penalizante y la solución regularizada obtenida mediante la minimización de funcionales TPG o TPGG cuando el parámetro de regularización es obtenido a partir de una regla heurística. En particular, se pretende obtener una cota para dicho error cuando el parámetro de regularización es elegido mediante la regla heurística presentada en el Capítulo 2.

Por simplicidad en las demostraciones de los resultados obtenidos, en la Sección 3.1 consideraremos el método que consiste en aproximar la solución del problema (3.1) mediante un

minimizante del funcional TPG- p dado por

$$\mathcal{J}_{\psi,\eta}^p(x) \doteq \frac{1}{p} \|\mathcal{T}x - y\|^p + \eta \psi(x), \quad (3.2)$$

donde $p \geq 1$ y $\psi : \mathcal{X} \rightarrow [0, +\infty]$ es un funcional convexo. Luego, en la Sección 3.2 consideraremos una versión aún más general de este método que consiste en aproximar la solución del problema (3.1) mediante un minimizante del funcional TPGG- p dado por

$$\mathcal{J}_{\tilde{\phi},\psi,\eta}^p(x) \doteq \frac{1}{p} \tilde{\phi}(\mathcal{T}x, y)^p + \eta \psi(x), \quad (3.3)$$

donde $p \geq 1$, $\tilde{\phi} : \mathcal{Y} \times \mathcal{Y} \rightarrow [0, +\infty]$ y $\psi : \mathcal{X} \rightarrow [0, +\infty]$ es un funcional convexo.

Es oportuno mencionar aquí que el funcional TP considerado en el Capítulo 1 es equivalente al funcional TPG-2 con $\psi(x) \doteq \frac{1}{2} \|x\|^2$. Además, el funcional TPG considerado en el Capítulo 1 resulta ser equivalente al funcional TPG-2 en (3.2), por lo que el funcional TPG- p extiende al funcional TPG. Por último, el funcional TPGG-1 en (3.3) coincide con el funcional TPGG presentado en el Capítulo 1, considerando $\phi(\mathcal{T}x, y) = \tilde{\phi}(\mathcal{T}x, y)$. Por lo tanto, el funcional TPGG- p es una generalización del funcional TPGG.

Finalmente, los resultados de existencia de solución regularizada y convergencia presentados en el Capítulo 1 siguen siendo válidos si consideramos el término de fidelidad dado por $\phi(\mathcal{T}x, y) \doteq \frac{1}{p} \tilde{\phi}(\mathcal{T}x, y)^p$ con $p \geq 1$. Más precisamente, es inmediato ver que si $\tilde{\phi}$ verifica las hipótesis III y V del Teorema 1.32 entonces ϕ también las verifica. Por otro lado, de la semicontinuidad inferior débil de $\tilde{\phi}$ y debido a que la función $f(t) = \frac{1}{p} t^p$ es no decreciente para $p \geq 1$, se sigue que ϕ verifica las hipótesis (G4) del Teorema 1.28 y IV del Teorema 1.32. En conclusión, los Teoremas 1.28 y 1.32 resultan válidos para el funcional TPGG- p .

3.1. Regularización de tipo Tikhonov-Phillips generalizado- p

Como hemos mencionado, a lo largo de esta sección consideraremos el método de regularización que consiste en aproximar la solución del problema (3.1) mediante un minimizante del funcional TPG- p dado por

$$\mathcal{J}_{\psi,\eta}^p(x) \doteq \frac{1}{p} \|\mathcal{T}x - y\|^p + \eta \psi(x), \quad (3.4)$$

donde $p \geq 1$ y $\psi : \mathcal{X} \rightarrow [0, +\infty]$ es un funcional convexo.

El objetivo principal de esta sección es hallar una estimación para el error a-posteriori, es decir, una cota para el error entre una solución ψ -minimizante de (3.1) y la solución regularizada x_η^δ , dada por un minimizante del funcional (3.4) para dato y^δ y para el caso en que el parámetro de regularización η sea elegido mediante una regla heurística. Como es usual en estos casos,

estimaremos este error utilizando la distancia de Bregman que es una forma natural de medir la desviación de elementos en un espacio de Banach con respecto a algún funcional convexo. Para definir la distancia de Bregman resulta necesario introducir los conceptos de subgradiente y subdiferencial de un funcional convexo, los cuales presentamos a continuación.

Definición 3.1. Sean \mathcal{X} un espacio de Banach, \mathcal{X}^* su espacio dual y $\psi : \mathcal{X} \rightarrow \mathbb{R} \cup \{\infty\}$ un funcional convexo. Decimos que $x^* \in \mathcal{X}^*$ es un *subgradiente* de ψ en x si $\psi(z) \geq \psi(x) + [x^*, z - x]$, $\forall z \in \mathcal{X}$, donde $[x^*, x]$ corresponde a la evaluación del funcional x^* en x , es decir $[x^*, x] \doteq x^*(x)$. El conjunto $\partial\psi(x)$ de todos los subgradientes de ψ en x se llama el *subdiferencial* de ψ en x .

Cabe mencionar aquí que el subgradiente es una generalización del concepto clásico de derivada para el caso de funcionales convexos ([55]).

Definición 3.2. Sean \mathcal{X} un espacio de Banach, $\psi : \mathcal{X} \rightarrow \mathbb{R} \cup \{\infty\}$ un funcional convexo, $x \in \mathcal{X}$ y $\xi \in \partial\psi(x)$. La *distancia de Bregman en x respecto de ξ y ψ* se define como

$$d_{\xi}^{\psi}(z, x) \doteq \psi(z) - \psi(x) - [\xi, z - x], \quad \forall z \in \mathcal{X}.$$

Es fácil ver que $d_{\xi}^{\psi}(x, x) = 0$ y que la convexidad del funcional ψ implica que $d_{\xi}^{\psi}(z, x) \geq 0$, para todo $z \in \mathcal{X}$. Más adelante veremos también que si \mathcal{X} es un espacio de Hilbert, la distancia de Bregman en un elemento x coincide con $\|x - z\|^2$, para todo $z \in \mathcal{X}$, cuando el funcional ψ es elegido como el cuadrado de la norma en dicho espacio. De esta manera, la distancia de Bregman extiende el concepto de norma.

Como es usual para problemas inversos mal condicionados, las estimaciones del error entre una solución regularizada y una solución ψ -minimizante se pueden obtener únicamente bajo supuestos adicionales de suavidad de la solución ψ -minimizante, conocidos como *condición fuente* (“source conditions”). En la Sección 3.2 veremos que esta condición fuente se encuentra estrechamente relacionada con la condición de optimalidad que verifica todo minimizante de un funcional convexo. En los resultados obtenidos en esta sección, la condición fuente estará vinculada al operador adjunto $\mathcal{T}^* : \mathcal{Y} \rightarrow \mathcal{X}^*$ de \mathcal{T} .

Como hemos visto en el Capítulo 1, si el parámetro de regularización es elegido mediante una regla a-priori, es posible probar la convergencia de las soluciones regularizadas obtenidas mediante la minimización de funcionales de tipo Tikhonov-Phillips generalizados a una solución de mínimo penalizante, cuando el nivel de ruido tiende a cero por derecha. Por otro lado, para reglas a-posteriori, como por ejemplo el principio de discrepancia de Morozov, existen resultados que establecen la convergencia de las soluciones regularizadas obtenidas a partir

de la minimización de funcionales de tipo Tikhonov-Phillips generalizados ([50]). Finalmente, para el caso de reglas heurísticas es posible encontrar estimaciones del error a-posteriori, es decir, cotas para el error entre una solución de mínimo penalizante y una solución regularizada obtenida mediante la minimización de funcionales de tipo Tikhonov-Phillips generalizados, para un nivel de ruido fijo.

A continuación, presentamos el primer resultado obtenido que establece una estimación para el error a-posteriori entre una solución ψ -minimizante del problema (3.1) y una solución regularizada x_η^δ , minimizante del funcional $\mathcal{J}_{\psi,\eta}^p$ con dato y^δ , cuando el parámetro de regularización es elegido mediante una regla heurística $\eta \doteq \eta(y^\delta)$. Este error se medirá utilizando la distancia de Bregman inducida por el penalizante ψ . De esta manera, dicha distancia controlará únicamente los errores que pueden ser distinguidos por el penalizante, mientras que los otros errores deben ser controlados por el término de fidelidad ([5]).

Teorema 3.3. *Sean \mathcal{X} un espacio de Banach, \mathcal{Y} es espacio de Hilbert, $\mathcal{T} \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$, $\psi : \mathcal{X} \rightarrow [0, +\infty]$, \tilde{x} una solución ψ -minimizante de (3.1) e $y^\delta \in \mathcal{Y}$ tal que $\|y - y^\delta\| \leq \delta$. Supongamos que existe $w \in \mathcal{Y}$ tal que $\xi = \mathcal{T}^*w \in \partial\psi(\tilde{x})$. Luego, para cada minimizante x_η^δ del funcional $\mathcal{J}_{\psi,\eta}^p$ dado en (3.4) con dato y^δ , se tiene que*

- si $p > 1$ y q es el conjugado de p , entonces

$$d_\xi^\psi(x_\eta^\delta, \tilde{x}) \leq \frac{\frac{\delta^p}{p} + \eta \delta \|w\| + \frac{1}{q}(\eta \|w\|)^q}{\eta};$$

- si $p = 1$ y $\eta \|w\| \leq 1$, entonces

$$d_\xi^\psi(x_\eta^\delta, \tilde{x}) \leq \delta \frac{1 + \eta \|w\|}{\eta}.$$

Demostración. Dado que x_η^δ es un minimizante del funcional $\mathcal{J}_{\psi,\eta}^p$ con dato y^δ , se sigue que

$$\begin{aligned} \frac{1}{p} \|\mathcal{T}x_\eta^\delta - y^\delta\|^p + \eta \psi(x_\eta^\delta) &\leq \frac{1}{p} \|\mathcal{T}\tilde{x} - y^\delta\|^p + \eta \psi(\tilde{x}) \\ &\leq \frac{\delta^p}{p} + \eta \psi(\tilde{x}), \end{aligned} \quad (\mathcal{T}\tilde{x} = y \text{ y } \|y - y^\delta\| \leq \delta)$$

y así

$$\frac{1}{p} \|\mathcal{T}x_\eta^\delta - y^\delta\|^p \leq \frac{\delta^p}{p} + \eta (\psi(\tilde{x}) - \psi(x_\eta^\delta)).$$

Luego,

$$\frac{1}{p} \|\mathcal{T}x_\eta^\delta - y^\delta\|^p + \eta d_\xi^\psi(x_\eta^\delta, \tilde{x})$$

$$\begin{aligned}
&\leq \frac{\delta^p}{p} + \eta \left(\psi(\tilde{x}) - \psi(x_\eta^\delta) + d_\xi^\psi(x_\eta^\delta, \tilde{x}) \right) \\
&= \frac{\delta^p}{p} - \eta [\xi, x_\eta^\delta - \tilde{x}] && \text{(por definición de distancia de Bregman)} \\
&= \frac{\delta^p}{p} + \eta [\mathcal{T}^* w, \tilde{x} - x_\eta^\delta] \\
&= \frac{\delta^p}{p} + \eta \langle w, \mathcal{T} \tilde{x} - \mathcal{T} x_\eta^\delta \rangle && (\mathcal{Y} \text{ espacio de Hilbert)} \\
&\leq \frac{\delta^p}{p} + \eta \|w\| \|\mathcal{T} x_\eta^\delta - y\| \\
&\leq \frac{\delta^p}{p} + \eta \|w\| \|\mathcal{T} x_\eta^\delta - y^\delta\| + \eta \|w\| \|y^\delta - y\| \\
&\leq \frac{\delta^p}{p} + \eta \|w\| \|\mathcal{T} x_\eta^\delta - y^\delta\| + \eta \|w\| \delta.
\end{aligned}$$

Finalmente,

$$\frac{1}{p} \|\mathcal{T} x_\eta^\delta - y^\delta\|^p - \eta \|w\| \|\mathcal{T} x_\eta^\delta - y^\delta\| + \eta d_\xi^\psi(x_\eta^\delta, \tilde{x}) \leq \frac{\delta^p}{p} + \eta \|w\| \delta. \quad (3.5)$$

- Si $p > 1$, por la desigualdad de Young tenemos que

$$\eta \|w\| \|\mathcal{T} x_\eta^\delta - y^\delta\| \leq \frac{1}{p} \|\mathcal{T} x_\eta^\delta - y^\delta\|^p + \frac{1}{q} (\eta \|w\|)^q,$$

lo cual junto a (3.5) permite deducir que $\eta d_\xi^\psi(x_\eta^\delta, \tilde{x}) - \frac{1}{q} (\eta \|w\|)^q \leq \frac{\delta^p}{p} + \eta \|w\| \delta$, de donde se sigue inmediatamente la desigualdad que queríamos probar.

- Si $p = 1$, de (3.5) se tiene que

$$(1 - \eta \|w\|) \|\mathcal{T} x_\eta^\delta - y^\delta\| + \eta d_\xi^\psi(x_\eta^\delta, \tilde{x}) \leq \delta + \eta \|w\| \delta,$$

y como por hipótesis $1 - \eta \|w\| \geq 0$ resulta $d_\xi^\psi(x_\eta^\delta, \tilde{x}) \leq \delta \frac{1 + \eta \|w\|}{\eta}$.

□

Observación 3.4. En el Capítulo 1 presentamos algunos términos de fidelidad diferentes del clásico y vimos en qué casos resulta conveniente su utilización. Uno de ellos era el término de fidelidad dado por $\|\mathcal{T}x - y\|_1$. Para este caso entonces debemos considerar $\tilde{\phi}(\mathcal{T}x, y) = \|\mathcal{T}x - y\|_1$ y $p = 1$ y, de esta manera, es posible obtener una estimación del error a-posteriori como consecuencia del Teorema 3.3.

Observación 3.5. Es oportuno destacar en el teorema anterior el caso $p = 2$ (correspondiente al término de fidelidad clásico) en el cual se obtiene la siguiente cota para el error a-posteriori

$$d_\xi^\psi(x_\eta^\delta, \tilde{x}) \leq \frac{\frac{\delta^2}{2} + \eta \delta \|w\| + \frac{1}{2} (\eta \|w\|)^2}{\eta} = \frac{1}{2} \left(\frac{\delta}{\sqrt{\eta}} + \sqrt{\eta} \|w\| \right)^2.$$

En [27] se presenta una cota para este error bajo los mismos supuestos del Teorema 3.3 para $p = 2$. Esta cota viene dada por

$$d_{\xi}^{\psi}(x_{\eta}^{\delta}, \tilde{x}) \leq \frac{\delta^2}{\eta} + \eta \|w\|^2.$$

Es sabido que si a y b son números reales no negativos entonces verifican la desigualdad $\frac{1}{2}(a + b)^2 \leq a^2 + b^2$, de donde se sigue que

$$\frac{1}{2} \left(\frac{\delta}{\sqrt{\eta}} + \sqrt{\eta} \|w\| \right)^2 \leq \frac{\delta^2}{\eta} + \eta \|w\|^2.$$

Es decir, la estimación obtenida en el Teorema 3.3 es inferior o igual a la obtenida en la prueba del Teorema 3.5 en [27].

Ahora, y como consecuencia del Teorema 3.3, presentamos una estimación para el error a-posteriori entre \tilde{x} y $x_{\eta^*}^{\delta}$, donde η^* es elegido a partir de la regla de elección del parámetro heurística presentada en el Capítulo 2, utilizando $\phi(\mathcal{T}x, y) \doteq \frac{1}{p} \|\mathcal{T}x - y\|^p$ con $p > 1$. Recordemos que para esta regla el parámetro de regularización óptimo η^* viene dado por

$$\eta^* = \alpha \frac{\|\mathcal{T}x_{\eta^*}^{\delta} - y^{\delta}\|^{p(1-d)}}{p^{1-d} (\psi(x_{\eta^*}^{\delta}) + \beta_0)},$$

con $0 < d < 1$ y $\alpha, \beta_0 \geq 0$. Esta cota se obtiene simplemente reemplazando el parámetro η por la regla η^* en el Teorema 3.3.

Corolario 3.6. Si $\eta^* = \alpha \frac{\|\mathcal{T}x_{\eta^*}^{\delta} - y^{\delta}\|^{p(1-d)}}{p^{1-d} (\psi(x_{\eta^*}^{\delta}) + \beta_0)}$, $p > 1$ y $\delta^* \doteq \|\mathcal{T}x_{\eta^*}^{\delta} - y^{\delta}\|$ entonces

$$d_{\xi}^{\psi}(x_{\eta^*}^{\delta}, \tilde{x}) \leq \left(\frac{\delta^{p-1} \psi(x_{\eta^*}^{\delta}) + \beta_0}{p^d \alpha (\delta^*)^{p(1-d)}} + \|w\| + \frac{1}{q} \frac{\alpha^{q-1}}{p^{(1-d)(q-1)} (\psi(x_{\eta^*}^{\delta}) + \beta_0)} \|w\|^q \right) \max \left\{ \delta, (\delta^*)^{q(1-d)} \right\}.$$

Observación 3.7. Es oportuno observar que del corolario anterior se obtiene la siguiente cota para el caso en que $p = 2$

$$d_{\xi}^{\psi}(x_{\eta^*}^{\delta}, \tilde{x}) \leq \left(\frac{\delta \psi(x_{\eta^*}^{\delta}) + \beta_0}{2^d \alpha (\delta^*)^{2(1-d)}} + \|w\| + \frac{\alpha}{2^{2-d} (\psi(x_{\eta^*}^{\delta}) + \beta_0)} \|w\|^2 \right) \max \left\{ \delta, (\delta^*)^{2(1-d)} \right\}. \quad (3.6)$$

Como consecuencia de la Observación 3.5 podemos decir entonces que la cota obtenida en (3.6) resulta inferior o igual a la cota presentada en [27] para esta regla.

3.2. Regularización de tipo Tikhonov-Phillips doblemente generalizado- p

En la sección anterior, hemos considerado el método que aproxima la solución del problema (3.1) mediante un minimizante de un funcional TPG- p , es decir, un funcional de Tikhonov-Phillips con término de fidelidad $\frac{1}{p}\|\mathcal{T}x - y\|^p$ y penalizante dado por un funcional convexo ψ . Para dicho método hemos encontrado una estimación del error a-posteriori cuando la regla de elección del parámetro es heurística y la hemos particularizado para el caso en que se utilice la regla presentada en el Capítulo 2. En esta sección, generalizaremos los resultados de cotas obtenidos en la Sección 3.1 al caso de funcionales TPGG- p , es decir, para funcionales de Tikhonov-Phillips con fidelidad y penalizante generalizados de la forma

$$\mathcal{J}_{\tilde{\phi}, \psi, \eta}^p(x) = \frac{1}{p} \tilde{\phi}(\mathcal{T}x, y)^p + \eta \psi(x), \quad (3.7)$$

donde $p \geq 1$, $\tilde{\phi} : \mathcal{Y} \times \mathcal{Y} \rightarrow [0, +\infty]$ y $\psi : \mathcal{X} \rightarrow [0, +\infty]$ es un funcional convexo.

En la sección anterior vimos que uno de los supuestos para encontrar estimaciones del error a-posteriori viene dado por una condición fuente y que la misma está relacionada con el operador adjunto \mathcal{T}^* de \mathcal{T} , para el caso en que \mathcal{Y} sea un espacio de Hilbert. En esta sección, donde supondremos que \mathcal{X} e \mathcal{Y} son espacios de Banach, la condición fuente estará estrechamente relacionada al operador *dual adjunto* $\mathcal{T}^\# : \mathcal{Y}^* \rightarrow \mathcal{X}^*$ de \mathcal{T} definido como el único operador lineal y acotado que verifica que $[\mathcal{T}^\# y^*, x] = [y^*, \mathcal{T}x]$, para todo $y^* \in \mathcal{Y}^*$ y $x \in \mathcal{X}$.

El siguiente teorema presenta una primera estimación para el error a-posteriori.

Teorema 3.8. *Sean \mathcal{X} e \mathcal{Y} espacios de Banach, $\mathcal{T} \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$, $\psi : \mathcal{X} \rightarrow [0, +\infty]$, $\tilde{\phi} : \mathcal{Y} \times \mathcal{Y} \rightarrow [0, +\infty]$, \tilde{x} una solución ψ -minimizante de (3.1) e $y^\delta \in \mathcal{Y}$ tal que $\tilde{\phi}(y, y^\delta) \leq \delta$. Supongamos que existe $w \in \mathcal{Y}^*$ tal que $\xi \doteq \mathcal{T}^\# w \in \partial\psi(\tilde{x})$. Luego, para cada minimizante x_η^δ del funcional $\mathcal{J}_{\tilde{\phi}, \psi, \eta}^p$ dado en (3.7) con dato y^δ se tiene que*

$$\frac{1}{p} \tilde{\phi}(\mathcal{T}x_\eta^\delta, y^\delta)^p + \eta d_\xi^\psi(x_\eta^\delta, \tilde{x}) \leq \frac{\delta^p}{p} + \eta [w, \mathcal{T}\tilde{x} - \mathcal{T}x_\eta^\delta].$$

Demostración. Dado que x_η^δ es un minimizante del funcional $\mathcal{J}_{\tilde{\phi}, \psi, \eta}^p$ para dato y^δ y $\tilde{\phi}(y, y^\delta) \leq \delta$, se sigue que

$$\frac{1}{p} \tilde{\phi}(\mathcal{T}x_\eta^\delta, y^\delta)^p + \eta \psi(x_\eta^\delta) \leq \frac{1}{p} \tilde{\phi}(\mathcal{T}\tilde{x}, y^\delta)^p + \eta \psi(\tilde{x}) \leq \frac{\delta^p}{p} + \eta \psi(\tilde{x}),$$

y, en consecuencia, $\frac{1}{p} \tilde{\phi}(\mathcal{T}x_\eta^\delta, y^\delta)^p \leq \frac{\delta^p}{p} + \eta (\psi(\tilde{x}) - \psi(x_\eta^\delta))$. Luego,

$$\frac{1}{p} \tilde{\phi}(\mathcal{T}x_\eta^\delta, y^\delta)^p + \eta d_\xi^\psi(x_\eta^\delta, \tilde{x}) \leq \frac{\delta^p}{p} + \eta (\psi(\tilde{x}) - \psi(x_\eta^\delta) + d_\xi^\psi(x_\eta^\delta, \tilde{x}))$$

$$\begin{aligned}
&= \frac{\delta^p}{p} + \eta \left[\mathcal{T}^\# w, \tilde{x} - x_\eta^\delta \right] \\
&= \frac{\delta^p}{p} + \eta \left[w, \mathcal{T} \tilde{x} - \mathcal{T} x_\eta^\delta \right],
\end{aligned}$$

donde la penúltima igualdad se sigue de la condición fuente y la última, de la definición del operador dual adjunto. \square

El siguiente teorema presenta una cota para el error a-posteriori que generaliza el resultado dado en el Teorema 3.3 y cuya prueba se basa en la primera estimación de dicho error obtenida en el Teorema 3.8.

Teorema 3.9. Sean \mathcal{X} e \mathcal{Y} espacios de Banach, $\mathcal{T} \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$, $\psi : \mathcal{X} \rightarrow [0, +\infty]$, $\tilde{\phi} : \mathcal{Y} \times \mathcal{Y} \rightarrow [0, +\infty]$, \tilde{x} una solución ψ -minimizante de (3.1) e $y^\delta \in \mathcal{Y}$ tal que $\tilde{\phi}(y, y^\delta) \leq \delta$ y $\tilde{\phi}(y^\delta, y) \leq \delta$. Supongamos que existe $w \in \mathcal{Y}^*$ tal que $\xi \doteq \mathcal{T}^\# w \in \partial\psi(\tilde{x})$ y que el funcional $\tilde{\phi}$ satisface las siguientes hipótesis:

- I. $\tilde{\phi}(y_1, y_3) \leq \tilde{\phi}(y_1, y_2) + \tilde{\phi}(y_2, y_3)$, para todo $y_1, y_2, y_3 \in \mathcal{Y}$;
- II. existe $C > 0$ tal que $\|y_1 - y_2\| \leq C \tilde{\phi}(y_2, y_1)$, $\forall y_1, y_2 \in \mathcal{R}(\mathcal{T})$.

Luego, para cada minimizante x_η^δ del funcional $\mathcal{J}_{\tilde{\phi}, \psi, \eta}^p$ dado en (3.7) con dato y^δ se tiene que

- si $p > 1$ y q es el conjugado de p entonces

$$d_\xi^\psi(x_\eta^\delta, \tilde{x}) \leq \frac{\frac{\delta^p}{p} + \eta \delta \|w\| C + \frac{1}{q}(\eta \|w\| C)^q}{\eta};$$

- si $p = 1$ y $\eta \|w\| C \leq 1$ entonces

$$d_\xi^\psi(x_\eta^\delta, \tilde{x}) \leq \delta \frac{1 + \eta \|w\| C}{\eta}.$$

Demostración. Sea x_η^δ un minimizante del funcional $\mathcal{J}_{\tilde{\phi}, \psi, \eta}^p$ con dato y^δ . Luego,

$$\begin{aligned}
&\frac{1}{p} \tilde{\phi}(\mathcal{T} x_\eta^\delta, y^\delta)^p + \eta d_\xi^\psi(x_\eta^\delta, \tilde{x}) \\
&\leq \frac{\delta^p}{p} + \eta [w, \mathcal{T} \tilde{x} - \mathcal{T} x_\eta^\delta] && \text{(por Teorema 3.8)} \\
&\leq \frac{\delta^p}{p} + \eta \|w\| \|\mathcal{T} \tilde{x} - \mathcal{T} x_\eta^\delta\| \\
&\leq \frac{\delta^p}{p} + \eta \|w\| C \tilde{\phi}(\mathcal{T} x_\eta^\delta, y^\delta) + \eta \|w\| C \tilde{\phi}(y^\delta, y) && \text{(por Hipótesis I y II y } \mathcal{T} \tilde{x} = y) \\
&\leq \frac{\delta^p}{p} + \eta \|w\| C \tilde{\phi}(\mathcal{T} x_\eta^\delta, y^\delta) + \eta \|w\| C \delta. && (\tilde{\phi}(y^\delta, y) \leq \delta)
\end{aligned}$$

Finalmente,

$$\frac{1}{p} \tilde{\phi}(\mathcal{T} x_\eta^\delta, y^\delta)^p - \eta \|w\| C \tilde{\phi}(\mathcal{T} x_\eta^\delta, y^\delta) + \eta d_\xi^\psi(x_\eta^\delta, \tilde{x}) \leq \frac{\delta^p}{p} + \eta \|w\| C \delta. \quad (3.8)$$

- Si $p > 1$, de la desigualdad de Young se sigue inmediatamente que

$$\eta \|w\| C \tilde{\phi}(\mathcal{T}x_\eta^\delta, y^\delta) \leq \frac{1}{p} \tilde{\phi}(\mathcal{T}x_\eta^\delta, y^\delta)^p + \frac{1}{q} (\eta \|w\| C)^q,$$

lo cual implica que

$$\eta d_\xi^\psi(x_\eta^\delta, \tilde{x}) - \frac{1}{q} (\eta \|w\| C)^q \leq \frac{1}{p} \tilde{\phi}(\mathcal{T}x_\eta^\delta, y^\delta)^p - \eta \|w\| C \tilde{\phi}(\mathcal{T}x_\eta^\delta, y^\delta) + \eta d_\xi^\psi(x_\eta^\delta, \tilde{x}) \leq \frac{\delta^p}{p} + \eta \|w\| C \delta,$$

donde la última desigualdad se sigue de (3.8). De esta manera, resulta

$$d_\xi^\psi(x_\eta^\delta, \tilde{x}) \leq \frac{\frac{\delta^p}{p} + \eta \delta \|w\| C + \frac{1}{q} (\eta \|w\| C)^q}{\eta}.$$

- Si $p = 1$, de (3.8) se tiene que $(1 - \eta \|w\| C) \tilde{\phi}(\mathcal{T}x_\eta^\delta, y^\delta) + \eta d_\xi^\psi(x_\eta^\delta, \tilde{x}) \leq \delta + \eta \|w\| C \delta$, y como por hipótesis $1 - \eta \|w\| C \geq 0$ se obtiene

$$d_\xi^\psi(x_\eta^\delta, \tilde{x}) \leq \delta \frac{1 + \eta \|w\| C}{\eta}.$$

□

Dado que toda norma satisface la desigualdad triangular y considerando $C = 1$, resulta inmediato que las hipótesis I y II del Teorema 3.9 se verifican cuando $\tilde{\phi}$ corresponde a la norma en el espacio \mathcal{Y} . De esta manera, si se considera el término de fidelidad $\frac{1}{p} \|\mathcal{T}x - y\|^p$, con $p \geq 1$, del Teorema 3.9 se obtienen las mismas estimaciones del error a-posteriori que en el Teorema 3.3. De aquí se sigue inmediatamente que el Teorema 3.9 constituye una generalización del Teorema 3.3.

Observación 3.10. Es oportuno mencionar aquí que en el Teorema 3.8 es necesario que se satisfaga una condición fuente y no se necesitan hipótesis adicionales sobre el término de fidelidad y el penalizante. De esta manera, si se consideran, por ejemplo, los términos de fidelidad $\frac{1}{2} \|\mathcal{T}x - y\|^2$, $\|\mathcal{T}x - y\|_1$ o $\text{KL}(\mathcal{T}x, y)$ es posible obtener una estimación del error a-posteriori para estos casos como consecuencia del Teorema 3.8. Por otro lado, es sabido que el funcional de Kullback-Leibler (1.14) no verifica la desigualdad triangular ([26]) correspondiente a la hipótesis I del Teorema 3.9. Sin embargo, como hemos mencionado, para los casos en que se considere la fidelidad $\phi(\mathcal{T}x, y) = \text{KL}(\mathcal{T}x, y)$ es posible obtener una estimación del error a-posteriori como consecuencia del Teorema 3.8.

Debido a que la distancia de Bregman en $\tilde{x} \in \mathcal{X}$ respecto del funcional convexo ψ involucra el término $[\xi, x - \tilde{x}]$, donde $\xi \in \partial\psi(\tilde{x})$ y $x \in \mathcal{X}$, en varios artículos se propone la utilización de desigualdades que contienen este término, denominadas *desigualdades variacionales*, y se prueba que son una herramienta poderosa para obtener tasas de convergencia ([18], [23], [49]).

Los primeros resultados de tasas de convergencia para minimizantes de funcionales de tipo Tikhonov-Phillips se basan en supuestos de suavidad de la solución respecto a un operador (en general no lineal) definido en un espacio de Hilbert ([17]) o de Banach ([5]). Dichos supuestos están expresados en términos de una condición fuente (generalmente asociada a una ecuación). Sin embargo, observaciones numéricas mostraron que, aún cuando los supuestos de suavidad no se verificaban, la tasa de convergencia no necesariamente se veía afectada significativamente. En 2007, B. Hofmann et al. ([23]) tuvieron en cuenta esta observación y debilitaron dichos supuestos reemplazando la condición fuente por una desigualdad variacional. Nos proponemos aquí utilizar desigualdades de este tipo para probar un resultado análogo al del Teorema 3.9 y así obtener estimaciones del error a-posteriori bajo este nuevo enfoque. En primer lugar, mostraremos la relación existente entre la condición fuente presente en el Teorema 3.9 y una desigualdad de tipo variacional.

Proposición 3.11. *Sean \mathcal{X} e \mathcal{Y} espacios de Banach, $\mathcal{T} \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$, $\psi : \mathcal{X} \rightarrow [0, +\infty]$, $\tilde{\phi} : \mathcal{Y} \times \mathcal{Y} \rightarrow [0, +\infty]$, \tilde{x} una solución ψ -minimizante de (3.1) tal que existe $w \in \mathcal{Y}^*$ con $\xi \doteq \mathcal{T}^\# w \in \partial\psi(\tilde{x})$. Supongamos además que*

$$\text{existe } C > 0 \text{ tal que } \|y_1 - y_2\| \leq C \tilde{\phi}(y_2, y_1), \forall y_1, y_2 \in \mathcal{R}(\mathcal{T}). \quad (3.9)$$

Entonces, existe $\beta \geq 0$ tal que se verifica la siguiente desigualdad variacional

$$[\xi, \tilde{x} - x] \leq \beta \tilde{\phi}(\mathcal{T}x, \mathcal{T}\tilde{x}), \forall x \in \mathcal{X}. \quad (3.10)$$

Demostración. Es fácil ver que, por hipótesis y por la definición de operador dual adjunto, para cada $x \in \mathcal{X}$ resulta que

$$[\xi, \tilde{x} - x] = [\mathcal{T}^\# w, \tilde{x} - x] = [w, \mathcal{T}\tilde{x} - \mathcal{T}x] \leq \|w\| \|\mathcal{T}\tilde{x} - \mathcal{T}x\| \leq \|w\| C \tilde{\phi}(\mathcal{T}x, \mathcal{T}\tilde{x}),$$

de donde se sigue que la desigualdad (3.10) se verifica tomando $\beta \doteq C \|w\|$. \square

En la Proposición 3.13 presentamos un resultado en algún sentido “recíproco” al dado en la proposición anterior. Para ello, haremos uso del siguiente resultado acerca del operador dual adjunto $\mathcal{T}^\#$ de \mathcal{T} ([49], Lema 8.21).

Lema 3.12. *Sean \mathcal{X} e \mathcal{Y} espacios normados, $\mathcal{T} \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ y $x^* \in \mathcal{X}^*$. Entonces, $x^* \in \text{Ran}(\mathcal{T}^\#)$ si y sólo si existe $\tilde{C} > 0$ tal que $[[x^*, x]] \leq \tilde{C} \|\mathcal{T}x\|$, $\forall x \in \mathcal{X}$.*

Proposición 3.13. Sean \mathcal{X} e \mathcal{Y} espacios de Banach, $\mathcal{T} \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$, $\psi : \mathcal{X} \rightarrow [0, +\infty]$, $\tilde{\phi} : \mathcal{Y} \times \mathcal{Y} \rightarrow [0, +\infty]$, \tilde{x} una solución ψ -minimizante de (3.1). Supongamos además que existen $\beta \geq 0$ y $\xi \in \partial\psi(\tilde{x})$ tales que $[\xi, \tilde{x} - x] \leq \beta \tilde{\phi}(\mathcal{T}x, \mathcal{T}\tilde{x})$, $\forall x \in \mathcal{X}$, y que

$$\text{existe } c > 0 \text{ tal que } c \tilde{\phi}(y_1, y_2) \leq \|y_1 - y_2\|, \forall y_1, y_2 \in \mathcal{R}(\mathcal{T}). \quad (3.11)$$

Entonces, existe $w \in \mathcal{Y}^*$ tal que $\xi = \mathcal{T}^\# w$.

Demostración. Por hipótesis tenemos que

$$[\xi, x] = [\xi, \tilde{x} - (\tilde{x} - x)] \leq \beta \tilde{\phi}(\mathcal{T}(\tilde{x} - x), \mathcal{T}\tilde{x}) \leq \beta \frac{1}{c} \|\mathcal{T}x\|,$$

para todo $x \in \mathcal{X}$. Luego, del Lema 3.12 concluimos que $\xi \in \text{Ran}(\mathcal{T}^\#)$, es decir existe $w \in \mathcal{Y}^*$ tal que $\xi = \mathcal{T}^\# w$. \square

Observación 3.14. En el caso particular en que $\tilde{\phi}(\mathcal{T}x, y) \doteq \|\mathcal{T}x - y\|$, de las Proposiciones 3.11 y 3.13 se obtiene inmediatamente la equivalencia entre la condición fuente y la desigualdad variacional (3.10). Esto es así ya que las hipótesis (3.9) y (3.11) se verifican considerando $c = C = 1$.

El siguiente teorema presenta un resultado análogo al del Teorema 3.9 con una modificación en sus hipótesis ya que se establecen estimaciones del error a-posteriori utilizando una desigualdad variacional en lugar de una condición fuente. Dada la Proposición 3.11 es inmediato ver que esta nueva hipótesis es más débil que las utilizadas en el Teorema 3.9.

Teorema 3.15. Sean \mathcal{X} e \mathcal{Y} espacios de Banach, $\mathcal{T} \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$, $\psi : \mathcal{X} \rightarrow [0, +\infty]$, $\tilde{\phi} : \mathcal{Y} \times \mathcal{Y} \rightarrow [0, +\infty]$, \tilde{x} una solución ψ -minimizante de (3.1) e $y^\delta \in \mathcal{Y}$ tal que $\tilde{\phi}(y, y^\delta) \leq \delta$ y $\tilde{\phi}(y^\delta, y) \leq \delta$. Supongamos que el funcional $\tilde{\phi}$ verifica las siguientes hipótesis:

- I. $\tilde{\phi}(y_1, y_3) \leq \tilde{\phi}(y_1, y_2) + \tilde{\phi}(y_2, y_3)$, para todo $y_1, y_2, y_3 \in \mathcal{Y}$;
- II. existen $\beta \in [0, \infty)$ y $\xi \in \partial\psi(\tilde{x})$ tal que $[\xi, \tilde{x} - x] \leq \beta \tilde{\phi}(\mathcal{T}x, \mathcal{T}\tilde{x})$, para todo $x \in \mathcal{X}$.

Luego, para cada minimizante x_η^δ del funcional $\mathcal{J}_{\tilde{\phi}, \psi, \eta}^p$ dado en (3.7) con dato y^δ se tiene que

- si $p > 1$ y q es el conjugado de p , entonces

$$d_\xi^\psi(x_\eta^\delta, \tilde{x}) \leq \frac{\frac{\delta^p}{p} + \eta \delta \beta + \frac{1}{q}(\eta \beta)^q}{\eta};$$

- si $p = 1$ y $\eta \beta \leq 1$, entonces

$$d_\xi^\psi(x_\eta^\delta, \tilde{x}) \leq \delta \frac{1 + \eta \beta}{\eta}.$$

Demostración. Análogamente a la prueba del Teorema 3.8, dado que x_η^δ es un minimizante del funcional $\mathcal{J}_{\tilde{\phi}, \psi, \eta}^p$ con dato y^δ y $\tilde{\phi}(y, y^\delta) \leq \delta$, se sigue inmediatamente que

$$\frac{1}{p} \tilde{\phi}(\mathcal{T}x_\eta^\delta, y^\delta)^p \leq \frac{\delta^p}{p} + \eta \left(\psi(\tilde{x}) - \psi(x_\eta^\delta) \right).$$

Luego,

$$\begin{aligned} & \frac{1}{p} \tilde{\phi}(\mathcal{T}x_\eta^\delta, y^\delta)^p + \eta d_\xi^\psi(x_\eta^\delta, \tilde{x}) \\ & \leq \frac{\delta^p}{p} + \eta \left(\psi(\tilde{x}) - \psi(x_\eta^\delta) + d_\xi^\psi(x_\eta^\delta, \tilde{x}) \right) \\ & = \frac{\delta^p}{p} + \eta [\xi, \tilde{x} - x_\eta^\delta] \\ & \leq \frac{\delta^p}{p} + \eta \beta \tilde{\phi}(\mathcal{T}x_\eta^\delta, y^\delta) + \eta \beta \tilde{\phi}(y^\delta, y) \quad (\text{por Hipótesis I y II y } \mathcal{T}\tilde{x} = y) \\ & \leq \frac{\delta^p}{p} + \eta \beta \tilde{\phi}(\mathcal{T}x_\eta^\delta, y^\delta) + \eta \beta \delta. \quad (\text{pues } \tilde{\phi}(y^\delta, y) \leq \delta) \end{aligned}$$

Finalmente,

$$\frac{1}{p} \tilde{\phi}(\mathcal{T}x_\eta^\delta, y^\delta)^p - \eta \beta \tilde{\phi}(\mathcal{T}x_\eta^\delta, y^\delta) + \eta d_\xi^\psi(x_\eta^\delta, \tilde{x}) \leq \frac{\delta^p}{p} + \eta \beta \delta,$$

y la prueba se sigue análogamente a la del Teorema 3.9 con β en lugar de C en 3.8. \square

De la demostración de la Proposición 3.11 se sigue que $\beta = C \|w\|$ por lo que las estimaciones del error a-posteriori obtenidas en los Teoremas 3.9 y 3.15 coinciden.

Es oportuno mencionar aquí que si consideramos el término de fidelidad dado por $\frac{1}{p} \|\mathcal{T}x - y\|^p$, es decir $\tilde{\phi}(\mathcal{T}x, y) = \|\mathcal{T}x - y\|$, la hipótesis I del Teorema 3.15 se satisface como consecuencia de la desigualdad triangular para la norma. De esta manera, a partir del Teorema 3.15, es posible obtener un resultado que establezca una estimación del error a-posteriori entre una solución de mínimo penalizante y una solución regularizada obtenida mediante la minimización de funcionales TPG- p . Nuevamente, es posible concluir que la estimación del error a-posteriori obtenida en el Corolario 3.16 coincide con la del Teorema 3.3 (en este caso, $\beta = \|w\|$ puesto que $C = 1$).

Corolario 3.16. Sean \mathcal{X} e \mathcal{Y} espacios de Banach, $\mathcal{T} \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$, $\psi : \mathcal{X} \rightarrow [0, +\infty]$, \tilde{x} una solución ψ -minimizante de (3.1) e $y^\delta \in \mathcal{Y}$ tal que $\|y - y^\delta\| \leq \delta$. Supongamos que existen $\beta \in [0, \infty)$ y $\xi \in \partial\psi(\tilde{x})$ tal que $[\xi, \tilde{x} - x] \leq \beta \|\mathcal{T}x - \mathcal{T}\tilde{x}\|$, para todo $x \in \mathcal{X}$. Luego, para cada minimizante x_η^δ del funcional $\mathcal{J}_{\psi, \eta}^p$ dado en (3.4) con dato y^δ se tiene que

- si $p > 1$ y q es el conjugado de p , entonces

$$d_{\xi}^{\psi}(x_{\eta}^{\delta}, \tilde{x}) \leq \frac{\frac{\delta^p}{p} + \eta \delta \beta + \frac{1}{q}(\eta \beta)^q}{\eta};$$

- si $p = 1$ y $\eta \beta \leq 1$, entonces

$$d_{\xi}^{\psi}(x_{\eta}^{\delta}, \tilde{x}) \leq \delta \frac{1 + \eta \beta}{\eta}.$$

Tal como hemos visto, los Teoremas 3.3 y 3.9 requieren de supuestos asociados al cumplimiento de una condición fuente. Veremos ahora que la misma se encuentra estrechamente relacionada con la condición de optimalidad que verifican los minimizantes de los funcionales TPG- p y TPGG- p . Para ello, serán necesarios algunos resultados y definiciones previos que veremos a continuación ([13], [50]).

En primer lugar, presentaremos el concepto de mapeo de dualidad que será relevante por su estrecha relación con el subdiferencial de una potencia de una norma. Para definir este mapeo, debemos introducir el concepto de *función peso* que consiste en una función continua y estrictamente creciente $f : [0, +\infty) \rightarrow [0, +\infty)$ que verifica $f(0) = 0$ y $\lim_{t \rightarrow \infty} f(t) = \infty$. A su vez, resulta necesario definir el conjunto $2^{\mathcal{X}^*}$ formado por todos los subconjuntos del espacio dual de \mathcal{X} , es decir, $2^{\mathcal{X}^*} \doteq \{E : E \subset \mathcal{X}^*\}$.

Definición 3.17. Sea \mathcal{X} un espacio de Banach. El *mapeo de dualidad* respecto de la función peso f es el mapeo $J : \mathcal{X} \rightarrow 2^{\mathcal{X}^*}$ dado por

$$J(x) = \{x^* \in \mathcal{X}^* : [x^*, x] = \|x^*\| \|x\|, \|x^*\| = f(\|x\|)\}.$$

A continuación, se presenta el mapeo de dualidad respecto de la función peso $f(t) = t^{p-1}$ con $p > 1$ para el caso de espacios de Hilbert.

Proposición 3.18. Sean \mathcal{X} un espacio de Hilbert y $f(t) = t^{p-1}$ con $p > 1$. El mapeo de dualidad J respecto de f está dado por $J(x) = \|x\|^{p-2} x$.

El siguiente teorema, denominado *Teorema de Asplund*, relaciona el subdiferencial de la primitiva de una función peso (la cual resulta convexa) y el mapeo de dualidad respecto de dicha función.

Teorema 3.19. Sean \mathcal{X} un espacio de Banach y f una función peso. Si $F(t) \doteq \int_0^t f(s) ds$ entonces el mapeo de dualidad respecto de f está dado por

$$J(x) = \partial F(\|x\|), \quad x \in \mathcal{X}.$$

El siguiente corolario resulta útil al momento de obtener el subdiferencial del funcional convexo $\frac{1}{p}\|x\|^p$ con $p > 1$ que se encuentra asociado al término de fidelidad considerado en la Sección 3.1.

Corolario 3.20. *Si \mathcal{X} es un espacio de Banach y $f(t) = t^{p-1}$ con $p > 1$, entonces el mapeo de dualidad J respecto de f está dado por $J(x) = \partial \left(\frac{1}{p}\|x\|^p \right)$.*

Utilizando los conceptos de distancia de Bregman y de mapeo de dualidad presentados anteriormente, es posible probar que la distancia de Bregman en un elemento x de un espacio de Hilbert coincide con $\|x - z\|^2$ para todo $z \in \mathcal{X}$ cuando el funcional ψ es elegido como el cuadrado de la norma en dicho espacio. En efecto, sean \mathcal{X} un espacio de Hilbert y $\psi : \mathcal{X} \rightarrow [0, +\infty)$ el funcional dado por $\psi(x) = \|x\|^2$. Luego, de la Proposición 3.18 y el Corolario 3.20 (con $f(t) = t$) se sigue inmediatamente que el subdiferencial de ψ en x tiene un único elemento, ξ , dado por $\xi \doteq 2J(x) = 2x$. Como consecuencia, obtenemos que

$$\begin{aligned} d_{\xi}^{\psi}(z, x) &= \psi(z) - \psi(x) - [\xi, z - x] \\ &= \|z\|^2 - \|x\|^2 - \langle 2x, z - x \rangle \\ &= \|z\|^2 - \|x\|^2 - 2\langle x, z \rangle + 2\|x\|^2 \\ &= \|z\|^2 + \|x\|^2 - 2\langle x, z \rangle \\ &= \|x - z\|^2, \end{aligned}$$

es decir, la distancia de Bregman en x coincide con $\|x - z\|^2$, para todo $z \in \mathcal{X}$.

Es bien sabido que un minimizante de una función diferenciable y convexa verifica la condición de optimalidad, es decir, la derivada de dicha función se anula en aquel minimizante. De manera análoga, un minimizante z de un funcional convexo \mathcal{J} definido sobre un espacio de Banach satisface la condición de optimalidad dada por $0 \in \partial\mathcal{J}(z)$, como se presenta en el siguiente resultado. Más aún, la condición de optimalidad implica la existencia de un minimizante.

Teorema 3.21. *Sean \mathcal{X} un espacio de Banach, $\mathcal{J} : \mathcal{X} \rightarrow \mathbb{R} \cup \{\infty\}$ un funcional convexo y $z \in \mathcal{D}(\mathcal{J})$ con $\mathcal{J}(z) < \infty$. Entonces, $\mathcal{J}(z) = \min_{x \in \mathcal{X}} \mathcal{J}(x)$ si y sólo si $0 \in \partial\mathcal{J}(z)$.*

Demostración. $\mathcal{J}(z) = \min_{x \in \mathcal{X}} \mathcal{J}(x) \Leftrightarrow \mathcal{J}(z) \leq \mathcal{J}(x) \forall x \in \mathcal{X} \Leftrightarrow 0 \in \partial\mathcal{J}(z)$. □

A continuación probaremos que, bajo determinadas hipótesis, la condición de optimalidad para el funcional $\mathcal{J}_{\psi, \eta}^2$, con $\psi(x) \doteq \frac{1}{2}\|x\|^2$ (dada por la existencia de minimizantes) implica la condición fuente presente en el enunciado del Teorema 3.3.

Proposición 3.22. Sean \mathcal{X} e \mathcal{Y} espacios de Hilbert, $\psi(x) \doteq \frac{1}{2}\|x\|^2$ y \tilde{x} una solución ψ -minimizante de (3.1). Supongamos que $\{x_\eta\}$ es una sucesión de minimizantes del funcional $\mathcal{J}_{\psi,\eta}^2$ tal que $x_\eta \xrightarrow{\omega} \tilde{x}$ cuando $\eta \rightarrow 0^+$ y que existe $w \in \mathcal{Y}$ tal que $-\frac{\mathcal{T}x_\eta - y}{\eta} \xrightarrow{\omega} w$. Entonces, $\mathcal{T}^*w \in \partial\psi(\tilde{x})$, donde $\mathcal{T}^* : \mathcal{Y} \rightarrow \mathcal{X}$ es el operador adjunto de \mathcal{T} .

Demostración. Dado que x_η es un minimizante del funcional $\mathcal{J}_{\psi,\eta}^2$, del Teorema 3.21 resulta que $0 \in \partial \left(\frac{1}{2}\|\mathcal{T} \cdot -y\|^2 + \eta \frac{1}{2}\|\cdot\|^2 \right) (x_\eta)$. Luego, por propiedad de la suma, producto por un escalar, traslación y composición del subgradiente ([55]) se tiene que

$$\begin{aligned} \partial \left(\frac{1}{2}\|\mathcal{T} \cdot -y\|^2 + \eta \frac{1}{2}\|\cdot\|^2 \right) (x_\eta) &= \partial \left(\frac{1}{2}\|\mathcal{T} \cdot -y\|^2 \right) (x_\eta) + \eta \partial \left(\frac{1}{2}\|\cdot\|^2 \right) (x_\eta) \\ &= \mathcal{T}^* \partial \left(\frac{1}{2}\|\cdot -y\|^2 \right) (\mathcal{T}x_\eta) + \eta \partial \left(\frac{1}{2}\|\cdot\|^2 \right) (x_\eta) \\ &= \mathcal{T}^* \partial \left(\frac{1}{2}\|\cdot\|^2 \right) (\mathcal{T}x_\eta - y) + \eta \partial \left(\frac{1}{2}\|\cdot\|^2 \right) (x_\eta). \end{aligned}$$

Entonces, $0 \in \mathcal{T}^* \left(\partial \frac{1}{2}\|\cdot\|^2 \right) (\mathcal{T}x_\eta - y) + \eta \partial \left(\frac{1}{2}\|\cdot\|^2 \right) (x_\eta)$ y podemos descomponer el funcional 0 como $0 = f_\eta + g_\eta$, donde $f_\eta \in \mathcal{T}^* \partial \left(\frac{1}{2}\|\cdot\|^2 \right) (\mathcal{T}x_\eta - y)$ y $g_\eta \in \eta \partial \left(\frac{1}{2}\|\cdot\|^2 \right) (x_\eta)$. Como \mathcal{X} e \mathcal{Y} son espacios de Hilbert, de la Proposición 3.18 y el Corolario 3.20, se sigue inmediatamente que $f_\eta = \mathcal{T}^* (\mathcal{T}x_\eta - y)$ y $g_\eta = \eta x_\eta$. Como $g_\eta = -f_\eta$ resulta que

$$x_\eta = \frac{-\mathcal{T}^* (\mathcal{T}x_\eta - y)}{\eta}. \tag{3.12}$$

y entonces, por hipótesis, $x_\eta \xrightarrow{\omega} \mathcal{T}^*w$. Por unicidad del límite concluimos que $\tilde{x} = \mathcal{T}^*w$. De la Proposición 3.18 y del Corolario 3.20 se tiene que $\partial\psi(\tilde{x}) = \partial \left(\frac{1}{2}\|\cdot\|^2 \right) (\tilde{x}) = \{\tilde{x}\}$, y así $\mathcal{T}^*w \in \partial\psi(\tilde{x})$. \square

De manera análoga al caso clásico, veremos ahora que la condición de optimalidad para el funcional $\mathcal{J}_{\psi,\eta}^p$, con $p > 1$ y penalizante ψ convexo y débilmente semicontinuo inferiormente implica la condición fuente presente en el enunciado del Teorema 3.3.

Proposición 3.23. Sean \mathcal{X} un espacio de Banach, \mathcal{Y} un espacio de Hilbert, $\psi : \mathcal{X} \rightarrow [0, +\infty]$ un funcional convexo y débilmente semicontinuo inferiormente y \tilde{x} una solución ψ -minimizante de (3.1). Supongamos que $\{x_\eta\}$ es una sucesión de minimizantes del funcional $\mathcal{J}_{\psi,\eta}^p$ con $p > 1$ tal que $x_\eta \xrightarrow{\omega} \tilde{x}$ cuando $\eta \rightarrow 0^+$ y que existe $w \in \mathcal{Y}$ tal que $-\|\mathcal{T}x_\eta - y\|^{p-2} \frac{\mathcal{T}x_\eta - y}{\eta} \xrightarrow{\omega} w$. Entonces, $\mathcal{T}^*w \in \partial\psi(\tilde{x})$, donde $\mathcal{T}^* : \mathcal{Y} \rightarrow \mathcal{X}^*$ es el operador adjunto de \mathcal{T} .

Demostración. Análogamente al caso clásico, como x_η minimiza el funcional $\mathcal{J}_{\psi,\eta}^p$ resulta del Teorema 3.21 que

$$0 \in \partial \left(\frac{1}{p} \|\mathcal{T} \cdot - y\|^p + \eta \psi(\cdot) \right) (x_\eta) = \mathcal{T}^* \partial \left(\frac{1}{p} \|\cdot\|^p \right) (\mathcal{T}x_\eta - y) + \eta \partial \psi(x_\eta)$$

y podemos descomponer el funcional 0 como $0 = f_\eta + g_\eta$, donde $f_\eta \in \mathcal{T}^* \partial \left(\frac{1}{p} \|\cdot\|^p \right) (\mathcal{T}x_\eta - y)$ y $g_\eta \in \eta \partial \psi(x_\eta)$. Como \mathcal{Y} es un espacio de Hilbert, de la Proposición 3.18 y el Corolario 3.20 se sigue que $f_\eta = \|\mathcal{T}x_\eta - y\|^{p-2} \mathcal{T}^* (\mathcal{T}x_\eta - y)$. Si denotamos $w_\eta \doteq -\frac{\|\mathcal{T}x_\eta - y\|^{p-2} (\mathcal{T}x_\eta - y)}{\eta}$ entonces $-\frac{f_\eta}{\eta} = \mathcal{T}^* w_\eta$ y debido a que $-f_\eta = g_\eta$ resulta que $\mathcal{T}^* w_\eta = -\frac{f_\eta}{\eta} = \frac{g_\eta}{\eta} \in \partial \psi(x_\eta)$.

Por hipótesis sabemos que $w_\eta \xrightarrow{\omega} w$ de donde se sigue que $\mathcal{T}^* w_\eta \xrightarrow{\omega} \mathcal{T}^* w$. Queremos ver que $\mathcal{T}^* w \in \partial \psi(\tilde{x})$. En efecto, como ψ es débilmente semicontinuo inferiormente se tiene que $\psi(\tilde{x}) \leq \liminf_{\eta \rightarrow 0^+} \psi(x_\eta)$. Por otra parte, puesto que $\mathcal{T}^* w_\eta \in \partial \psi(x_\eta)$, se obtiene la siguiente desigualdad $\psi(z) \geq \psi(x_\eta) + [\mathcal{T}^* w_\eta, z - x_\eta]$, $\forall z \in \mathcal{X}$.

Tomando límite inferior resulta

$$\psi(z) \geq \liminf_{\eta \rightarrow 0^+} \psi(x_\eta) + \lim_{\eta \rightarrow 0^+} [\mathcal{T}^* w_\eta, z - x_\eta] \geq \psi(\tilde{x}) + [\mathcal{T}^* w, z - \tilde{x}],$$

para todo $z \in \mathcal{X}$, de donde se sigue que $\mathcal{T}^* w \in \partial \psi(\tilde{x})$. \square

Finalmente, veremos que la condición de optimalidad asociada al funcional TPGG- p con $p > 1$ y penalizante convexo y débilmente semicontinuo inferiormente implica la condición fuente presente en el enunciado del Teorema 3.9.

Proposición 3.24. Sean \mathcal{X} e \mathcal{Y} espacios de Banach, $\psi : \mathcal{X} \rightarrow [0, +\infty]$ un funcional convexo y débilmente semicontinuo inferiormente y \tilde{x} una solución ψ -minimizante de (3.1). Supongamos que $\{x_\eta\}$ es una sucesión de minimizantes del funcional $\mathcal{J}_{\phi,\psi,\eta}^p$ con $p > 1$ tal que $x_\eta \xrightarrow{\omega} \tilde{x}$ cuando $\eta \rightarrow 0^+$ y que

$$\text{para cada } \eta > 0, \text{ existe } w_\eta \in \mathcal{Y}^* \text{ tal que } \partial \left(-\frac{1}{\eta} \frac{\tilde{\phi}(\mathcal{T} \cdot, y)^p}{p} \right) (x_\eta) = \{w_\eta\}. \quad (3.13)$$

Si existe $w \in \mathcal{Y}^*$ tal que $w_\eta \xrightarrow{\omega} w$, entonces $\mathcal{T}^\# w \in \partial \psi(\tilde{x})$, donde $\mathcal{T}^\# : \mathcal{Y}^* \rightarrow \mathcal{X}^*$ es el operador dual adjunto de \mathcal{T} .

Demostración. Dado que x_η minimiza el funcional $\mathcal{J}_{\phi,\psi,\eta}^p$, por Teorema 3.21 se sigue inmediatamente que $0 \in \partial \left(\frac{1}{p} \tilde{\phi}(\mathcal{T} \cdot, y)^p + \eta \psi(\cdot) \right) (x_\eta)$. Luego, por propiedad de la suma, producto por un escalar y composición del subgradiente ([55]) se tiene que

$$0 \in \mathcal{T}^\# \partial \left(\frac{1}{p} \tilde{\phi}(\mathcal{T} \cdot, y)^p \right) (x_\eta) + \eta \partial \psi(x_\eta)$$

y entonces podemos descomponer el funcional nulo como $0 = f_\eta + g_\eta$, con

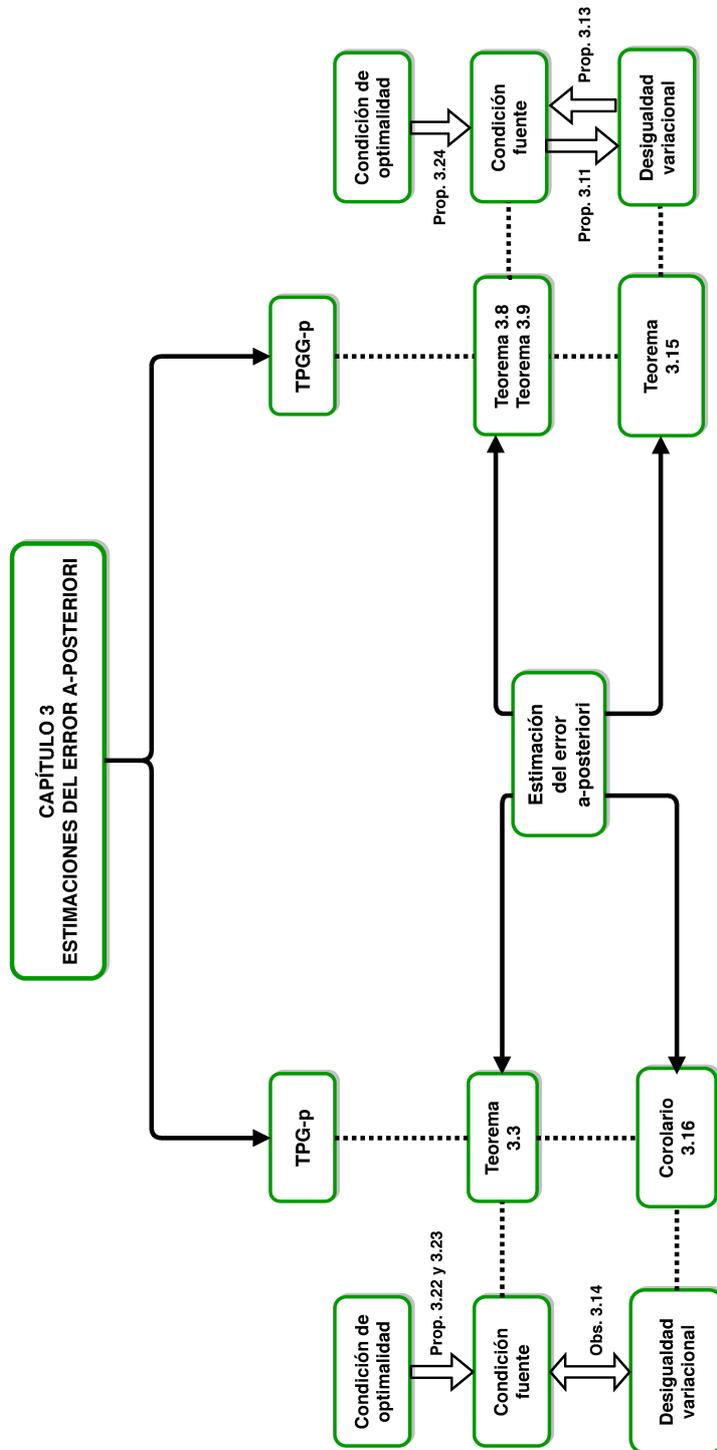
$$f_\eta \in \mathcal{T}^\# \partial \left(\frac{1}{p} \tilde{\phi}(\mathcal{T}\cdot, y)^p \right) (x_\eta) \text{ y } g_\eta \in \eta \partial \psi(x_\eta).$$

Ahora, por hipótesis sabemos que $\partial \left(-\frac{1}{\eta} \frac{\tilde{\phi}(\mathcal{T}\cdot, y)^p}{p} \right) (x_\eta) = \{w_\eta\}$ de donde resulta que $-\frac{f_\eta}{\eta} = \mathcal{T}^\# w_\eta$ y, como $g_\eta = -f_\eta$, se tiene que $\mathcal{T}^\# w_\eta = \frac{g_\eta}{\eta} \in \partial \psi(x_\eta)$.

Análogamente a la prueba de la Proposición 3.23, como consecuencia de la semicontinuidad inferior débil de ψ y debido a que $w_\eta \xrightarrow{\omega} w$ y $x_\eta \xrightarrow{\omega} \tilde{x}$ cuando $\eta \rightarrow 0^+$, se prueba que $\mathcal{T}^\# w \in \partial \psi(\tilde{x})$. \square

Observación 3.25. Es oportuno mencionar que, bajo ciertas hipótesis adicionales, el supuesto (3.13) realizado sobre el subdiferencial de $\frac{1}{p} \tilde{\phi}(\mathcal{T}\cdot, y)^p$ con $p > 1$ se verifica. Por ejemplo, si $\tilde{\phi}$ corresponde a la norma en un espacio de Hilbert, de la Proposición 3.18 se tiene que el subdiferencial $\partial \left(\frac{1}{p} \tilde{\phi}(\mathcal{T}\cdot, y)^p \right) (x_\eta)$ posee un único elemento. Por otro lado, si $\tilde{\phi}$ es Gateaux diferenciable entonces $\partial \tilde{\phi}(x) = \{\nabla \tilde{\phi}(x)\}$ ([50]).

En conclusión, a lo largo de este capítulo hemos obtenido estimaciones del error a-posteriori entre una solución de mínimo penalizante y una solución regularizada obtenida a partir de la minimización de los funcionales de Tikhonov-Phillips generalizados TPG- p y TPGG- p , para nivel de ruido fijo y para el caso en que se utilice una regla de elección de parámetro heurística.



Conclusiones y trabajos futuros

En el Capítulo 1 de esta tesis presentamos las herramientas matemáticas determinísticas clásicas en la teoría de Problemas Inversos Mal Condicionados. El abordaje de estos problemas requiere de diversas técnicas de regularización. En particular, estudiamos el tradicional método de Tikhonov-Phillips y propusimos dos versiones generalizadas del mismo: con penalizante generalizado y doblemente generalizado, es decir, con fidelidad y penalizante generales. En ese marco, mostramos cuáles son las principales ventajas de utilizar estas versiones generalizadas y probamos que estos métodos constituyen regularizaciones en espacios de Banach. En ese sentido se obtuvieron resultados de existencia y unicidad de solución; existencia de solución de mínimo penalizante y un resultado de convergencia asociado a una regla de elección del parámetro a-priori.

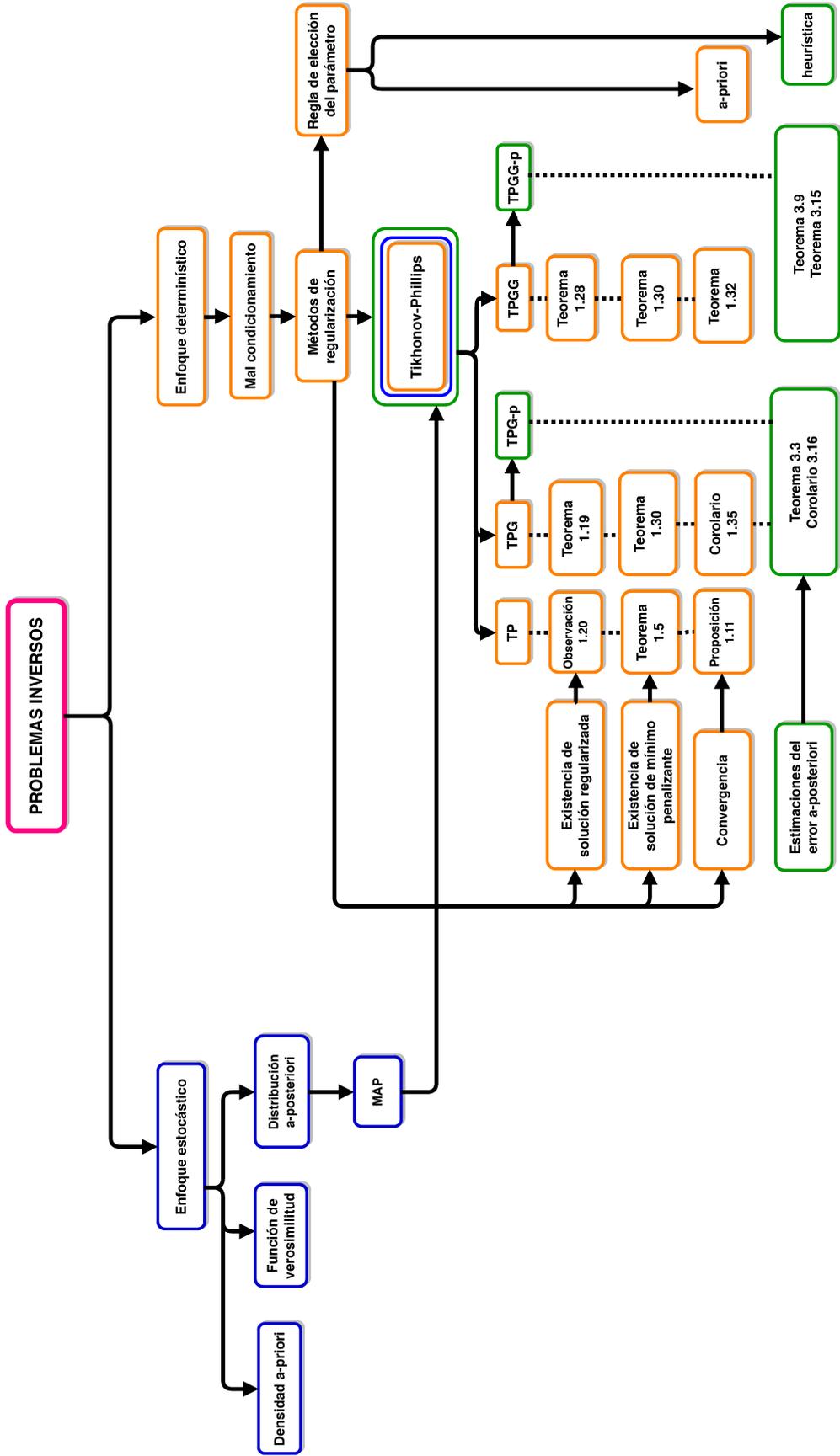
En el Capítulo 2 estudiamos el enfoque estocástico para la resolución de un problema inverso mal condicionado. Vimos allí que una característica que distingue a los enfoques determinístico y estocástico, es que en este último todas las variables involucradas son consideradas aleatorias dando como solución del problema inverso una distribución de probabilidad. A pesar de que esta manera de resolver el problema parece completamente desconectada de los métodos clásicos, dejamos en evidencia que existe un fuerte vínculo entre las versiones determinística y estocástica de resolución. Más precisamente, mostramos que las distribuciones asociadas al ruido y a la variable de interés están fuertemente relacionadas a los términos de fidelidad y penalizantes, respectivamente. Más aún, cuando las distribuciones asociadas a dichas variables pertenecen a una familia exponencial, mostramos que el estimador máximo a-posteriori coincide con la solución obtenida mediante la minimización de un funcional de tipo Tikhonov-Phillips doblemente generalizado. En este mismo capítulo, y en relación con los modelos jerárquicos, presentamos una regla de elección del parámetro heurística que fuera propuesta por K. Ito, B. Jin y J. Zou en [27]. Con el principal objetivo de mostrar el desempeño de la regla y el uso adecuado de diferentes penalizantes, se presentan aplicaciones en procesamiento de señales e imágenes.

En el Capítulo 3, presentamos los resultados de estimaciones del error a-posteriori obtenidos para el caso en que la regla de elección del parámetro es heurística. Es decir, se obtienen cotas para el error entre una solución de mínimo penalizante y una solución regularizada obtenida a partir de la minimización de funcionales de tipo Tikhonov-Phillips doblemente generalizados. Luego, estos resultados se particularizan a la regla construida en [27] y mostramos que la cota presentada por los autores en dicho trabajo resulta ser mayor o igual que la obtenida en este trabajo de tesis.

Los primeros resultados obtenidos en este capítulo, se basan principalmente en un supuesto dado por una condición fuente. Sin embargo, vimos que es posible reemplazar este supuesto por una desigualdad variacional obteniendo resultados análogos de estimaciones del error a-posteriori. Finalmente, se establecieron relaciones entre la condición de optimalidad asociada al problema, la condición fuente y la desigualdad variacional.

Para dar continuidad al trabajo de investigación desarrollado en esta tesis, nos proponemos avanzar en el estudio de los siguientes temas:

- Convergencia del método de regularización de Tikhonov-Phillips doblemente generalizado para el caso en el que el parámetro de regularización sea elegido mediante una regla a-posteriori.
- Generalización de la regla heurística estudiada en el contexto de familias exponenciales generales.
- Vínculo entre la elección de las funciones asociadas a las familias exponenciales y la regularización mediante un funcional de Tikhonov-Phillips con penalizante múltiple.
- Reglas de elección del parámetro múltiples, asociadas a funcionales de tipo Tikhonov-Phillips doblemente generalizados con penalizante múltiple.
- Generalización de los resultados obtenidos de estimaciones del error a-posteriori al caso de regularización mediante funcionales de tipo Tikhonov-Phillips doblemente generalizados con penalizante múltiple.



Bibliografía

1. R. Acar and C. R. Vogel, *Analysis of bounded variation penalty methods for ill-posed problems*, Inverse Problems **10** (1994), 1217–1229.
2. A. B. Bakushinskii, *Remarks on choosing a regularization parameter using the quasi-optimality and ratio criterion*, USSR Comp. Math. Math. Phys. **24** (1984), 181–182.
3. J. Besag, *On the statistical analysis of dirty pictures (with discussions)*, J.R. Stat. Soc. **B48** (1986), 259–302.
4. ———, *Towards bayesian image analysis*, J. Appl. Stat. **16** (1989), 395–406.
5. M. Burger and S. Osher, *Convergence rates of convex variational regularization*, Inverse Problems **20** (2004), no. 5, 1411–1421.
6. D. Calvetti, F. Sgallari, and E. Somersalo, *Image inpainting with structural bootstrap priors*, Image and Vision Computing **24** (2006), 782–793.
7. D. Calvetti and E. Somersalo, *A Gaussian hypermodel to recover blocky objects*, Inverse Problems **23** (2007), no. 2, 733–754.
8. ———, *Hypermodels in the Bayesian imaging framework*, Inverse Problems **24** (2008), no. 3, 034013, 20.
9. ———, *Large-scale inverse problems and quantification of uncertainty*, ch. Subjective knowledge or objective belief? An oblique look to Bayesian methods, pp. 33–70, Wiley, 2010.
10. A. Chambolle and J. L. Lions, *Image recovery via total variation minimization and related problems*, Numer. Math. **76** (1997), 167–188.
11. T. F. Chan and S. Esedoglu, *Aspects of total variation regularized L^1 function approximation*, SIAM Journal on Applied Mathematics **65** (2005), no. 5, 1817–1837.
12. G. Chavent and K. Kunisch, *Regularization of linear least squares problems by total bounded variation*, ESAIM: Control, Optimisation, and Calculus of Variations **2** (1997), 359–376.
13. I. Cioranescu, *Geometry of Banach spaces, duality mappings and nonlinear problems*, vol. 62, Springer Science & Business Media, 2012.
14. D. Dobson and F. Santosa, *An image enhancement technique for electrical impedance tomography*, Inverse Problems **10** (1994), 317–334.
15. P. B. Eggermont, *Maximum entropy regularization for Fredholm integral equations of the first kind*, SIAM J. Math. Anal. **24** (1993), 1557–1576.
16. H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of inverse problems*, Mathematics and its Applications, vol. 375, Kluwer Academic Publishers Group, Dordrecht, 1996.
17. H.W. Engl, K. Kunisch, and A. Neubauer, *Convergence rates for Tikhonov regularization of non-linear ill-posed problems*, Inverse Problems **5** (1989), no. 4, 523–540.

18. J. Flemming and B. Hofmann, *A new approach to source conditions in regularization with general residual term*, Numerical Functional Analysis and Optimization **31** (2010), 254–284.
19. S. Geman and D. Geman, *Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images*, IEEE Trans. Pattern Anal. Mach. Intell. **6** (1984), 721–741.
20. G. H Golub, M. T. Heath, and G. Wahbam, *Generalized cross-validation as a method for choosing a good ridge parameter*, Technometrics **21** (1979), 215–223.
21. P. C. Hansen, J. G. Nagy, and D. P. O’Leary, *Deblurring images: matrices, spectra, and filtering*, Fundamentals of Algorithms, vol. 3, SIAM, Philadelphia, 2006.
22. P.C. Hansen, *Analysis of discrete ill-posed problems by means of the L-curve*, SIAM Review **34**, no. 4 (1992), 561–580.
23. B. Hofmann, B. Kaltenbacher, C. Pöschl, and O. Scherzer, *A convergence rates result for Tikhonov regularization in Banach spaces with non-smooth operators*, Inverse Problems **23** (2007), no. 3, 987.
24. F. Ibarrola, G. Mazziери, R. Spies, and K. Temperini, *Anisotropic $BV - L^2$ regularization of linear inverse ill-posed problems*, Journal of Mathematical Analysis and Applications **450** (2017), 427–443.
25. B. Ito, K.; Jin and T. Takeuchi, *A regularization parameter for nonsmooth Tikhonov regularization*, SIAM Journal on Scientific Computing **33** (2011), 1415–1438.
26. K. Ito and B. Jin, *Inverse problems: Tikhonov theory and algorithms*, Series on applied mathematics, vol. 22, World Scientific, 2015.
27. K. Ito, B. Jin, and J. Zou, *A new choice rule for regularization parameters in Tikhonov regularization*, Applicable Analysis **90** (2011), 1521–1544.
28. B. Jin and J. Zou, *Augmented Tikhonov regularization*, Inverse Problems **25** (2008), no. 2, 025001.
29. J. Kaipio and E. Somersalo, *Statistical and computational inverse problems*, Applied Mathematical Sciences, vol. 160, Springer-Verlag, New York, 2005.
30. J. B. Keller, *Inverse problems*, Amer. Math. Monthly **83** (1976), no. 2, 107–118.
31. R. L. Lagendijk and J. Biemond, *Iterative identification and restoration of images*, Kluwer, Boston, Dordrecht, London, 1991.
32. P. Lamm, *Some recent developments and open problems in solution methods for mathematical inverse problems*, Proceedings XXIV National Congress of Applied and Computational Mathematics (2001).
33. E. L. Lehmann, *Theory of point estimation*, Wiley series in probability and mathematical statistics: Probability and mathematical statistics, Wiley, 1983.
34. G. L. Mazziери, R. D. Spies, and K. G. Temperini, *Existence, uniqueness and stability of minimizers of generalized Tikhonov-Phillips functionals*, Journal of Mathematical Analysis and Applications **396** (2012), 396–411.
35. ———, *Directional convergence of spectral regularization method associated to families of closed operators*, Computational and Applied Mathematics **32** (2013), 119–134.
36. ———, *Mixed spatially varying $L^2 - BV$ regularization of inverse ill-posed problems*, Journal of Inverse and Ill-Posed Problems **23** (2015), 571–585.
37. M. Z. Nashed, *Inner, outer, and generalized inverses in Banach and Hilbert spaces*, Numer. Funct. Anal. Optim. **9(3-4)** (1987), 261–325.

-
38. F. Natterer, *The mathematics of computerized tomography*, B. G. Teubner, Stuttgart, 1986.
 39. F. Natterer and F. Wübbeling, *Mathematical methods in image reconstruction*, SIAM, 2001.
 40. B. Nilsson, M. Johansson, A. Heyden, S. Nelander, and T. Fioretos, *An improved method for detecting and delineating genomic regions with altered gene expression in cancer*, *Genome Biol.* **9** (2008), no. 1. MR doi: 10.1186/gb-2008-9-1-r13
 41. S. Osher, A. Sole, and L. Vese, *Image decomposition and restoration using total variation minimization and the H^{-1} norm*, *SIAM Journal on Multiscale Modeling and Simulation* **1:3** (2003), 349–370.
 42. S. Osher and L. Vese, *Modeling textures with total variation minimization and oscillating patterns in image processing*, *Journal of Scientific Computing* **19:1-3** (2003), 553–572.
 43. D. L. Phillips, *A technique for the numerical solution of certain integral equations of the first kind*, *J. Assoc. Comput. Mach.* **9** (1962), 84–97.
 44. C. Pöschl, *Tikhonov regularization with general residual term*, PhD thesis (Dissertationsschrift). University of Innsbruck. Austria (2008).
 45. E. Resmerita and R. S. Anderssen, *Joint additive Kullback–Leibler residual minimization and regularization for linear inverse problems*, *Mathematical Methods in the Applied Sciences* **30** (2007), no. 13, 1527–1544.
 46. S.M. Ross, *Simulación*, Prentice Hall, México, 1999.
 47. L. I. Rudin, S. Osher, and Fatemi E., *Nonlinear total variation based noise removal algorithms (proceedings of the 11th annual international conference of the center for nonlinear studies)*, *Physica D* **60** (1992), 259–268.
 48. L. I. Rudin, S. Osher, and C. Fu, *Total variation based restoration of noisy, blurred images*, Preprint, Cognitech, Inc., 2800-28thth St., Suite 101, Santa Monica, CA 90405 (1994).
 49. O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen, *Variational methods in imaging*, *Applied Mathematical Sciences*, vol. 167, Springer, New York, 2009.
 50. T. Schuster, B. Kaltenbacher, B. Hofmann, and K. Kazimierski, *Regularization methods in Banach spaces*, de Gruyter, Berlin, New York, 2012.
 51. D. Strong and T. F. Chan, *Edge-preserving and scale-dependent properties of total variation regularization*, *Inverse Problems* **19** (2003), no. 6, S165–S187.
 52. A. N. Tikhonov, *Regularization of incorrectly posed problems*, *Soviet Math. Dokl.* **4** (1963), 1624–1627.
 53. ———, *Solution of incorrectly formulated problems and the regularization method*, *Soviet Math. Dokl.* **4** (1963), 1035–1038.
 54. C. R. Vogel and M. Oman, *Iterative methods for total variation denoising*, *Siam J. Sci. Computing* **17** (1996), 227–238.
 55. E. Zeidler, *Nonlinear functional analysis and its applications*, *Nonlinear functional analysis and its applications*, no. v. 1, Springer, 1986.