

## 19. Matemática Aplicada

### **Funcionales de tipo Tikhonov-Phillips generalizado: regla de elección de parámetro y cota de error a-posteriori.**

Lic. Carrió, María J.<sup>1</sup>; Dra. Mazzieri, Gisela L.<sup>2</sup>; Dra. Temperini, Karina G.<sup>2</sup>

<sup>1</sup> autor; <sup>2</sup> orientador

josefina.carrio@gmail.com; glmazzieri@santafe-conicet.gov.ar; ktemperini@santafe-conicet.gov.ar

Instituto de Matemática Aplicada del Litoral (CONICET-UNL)

Facultad de Bioquímica y Ciencias Biológicas - Universidad Nacional del Litoral

### **Resumen**

En este trabajo se estudia una regla de elección de parámetro de regularización para un funcional de Tikhonov-Phillips generalizado. A su vez, se obtiene una cota de error a-posteriori entre la solución exacta del problema y la solución regularizada obtenida, considerando como término de fidelidad  $\frac{1}{p} \|Kx - y\|^p$  y como penalizante un funcional convexo  $\psi(x)$ . Finalmente, se presenta un algoritmo para la implementación de la regla de elección del parámetro de regularización y se lo aplica a ejemplos numéricos de procesamiento de señales e imágenes.

**Palabras clave:** Problema inverso, regularización, Tikhonov-Phillips, regla de elección de parámetro, cota de error a-posteriori.

### **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 ha sido en buena parte consecuencia de aplicaciones en otras ciencias y en la industria. Desde el punto de vista de dichas aplicaciones existen dos motivaciones principales para el estudio de dichos problemas. Primero, la necesidad de determinar estados pasados

o parámetros de un sistema físico y segundo, la necesidad de saber cómo se debe influenciar un sistema a través de su estado presente o sus parámetros, con el objeto de llevarlo a un cierto estado deseado en el futuro. Podría decirse entonces que los problemas inversos tratan de determinar las causas de un efecto observado o deseado.

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

$$\mathcal{K}x = y, \quad (1)$$

donde  $\mathcal{K}$  es un operador lineal y acotado entre los espacios normados  $\mathcal{X}$  e  $\mathcal{Y}$  (en general,  $\mathcal{X}$  e  $\mathcal{Y}$  son espacios de Hilbert), e  $y$  es el dato que se supone conocido, quizás con algún nivel de error. Cuando la inversa generalizada de Moore-Penrose de  $\mathcal{K}$ ,  $\mathcal{K}^\dagger$ , no es acotada, entonces se dice que el problema es *mal condicionado* en el sentido de Hadamard, y pequeños errores en  $y$  pueden resultar en errores arbitrariamente grandes en  $x$  (Engl, Hanke & Neubauer, 1996). Esta falta de dependencia continua de los datos se traduce en inestabilidad desde el punto de vista de la resolución numérica del problema. 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. En el contexto del estudio de problemas inversos, a tales herramientas y métodos se los denomina *métodos de regularización*.

Problemas de este tipo se presentan en diversas áreas y en una gran variedad de aplicaciones, por ejemplo, en el tratamiento de imágenes y señales (Lagendijk & Biemond, 1991). En estos problemas, 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.

Uno de los métodos de regularización más conocido, y sin dudas el más utilizado, es el método de Tikhonov-Phillips, que fue propuesto independientemente por Phillips y Tikhonov en 1962 y 1963, respectivamente (Phillips, 1962; Tikhonov, 1963a, 1963b). Este método propone aproximar la solución del problema inverso (1) mediante el minimizante  $x_\eta$  del funcional

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

donde  $\eta$  es una constante positiva conocida como *parámetro de regularización*, el término  $\|\mathcal{K}x - y\|^2$  se conoce como *término de fidelidad* y  $\|x\|^2$  como *funcional de penalización*.

En este trabajo consideraremos el funcional de Tikhonov-Phillips generalizado

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

donde  $\phi: \mathcal{Y} \times \mathcal{Y} \rightarrow [0, \infty]$  y  $\psi: \mathcal{X} \rightarrow [0, \infty]$ .

## Objetivos

El presente trabajo tiene como principal objetivo el estudio de una regla de elección de parámetro de regularización propuesta en Ito, Jin y Zou (2011) para un funcional de Tikhonov-Phillips generalizado. Asimismo, obtendremos una cota para el error a-posteriori, es decir, una cota para el error entre la solución exacta del problema (1) y la solución regularizada obtenida a partir de la regla propuesta,

Universidad Nacional de Cuyo | Mendoza | Argentina  
utilizando como término de fidelidad  $\phi(\mathcal{K}x, y) := \frac{1}{p} \|\mathcal{K}x - y\|^p$  y como penalizante un funcional convexo  $\psi$ . Esta cota mejora la obtenida en Ito et al., para el caso  $p = 2$ . Por último, se presentará un algoritmo numérico para la implementación de esta regla y se lo aplicará a ejemplos numéricos de procesamiento de señales e imágenes.

### Materiales y Métodos

La metodología de trabajo consistió esencialmente en el estudio individual, búsqueda bibliográfica y redacción de los resultados que se obtuvieron. Para la implementación numérica de algunos resultados se utilizó el software Matlab R2017b.

### Resultados y Discusión

A lo largo de este trabajo consideraremos el problema inverso de determinar la incógnita  $x \in \mathcal{X}$  en la ecuación

$$\mathcal{K}x = y^\delta, \quad (2)$$

donde  $\mathcal{K}: \mathcal{X} \rightarrow \mathcal{Y}$  es un operador lineal y acotado entre los espacios de Banach  $\mathcal{X}$  e  $\mathcal{Y}$ , con  $\mathcal{X}$  reflexivo,  $y^\delta \in \mathcal{Y}$  es una versión con ruido del dato exacto  $y = \mathcal{K}x^\dagger$  y  $x^\dagger$  corresponde a la solución exacta.

Como hemos mencionado, para resolver el problema (2) consideraremos el funcional de Tikhonov-Phillips generalizado

$$\mathcal{J}_{\eta, y^\delta}(x) = \phi(\mathcal{K}x, y^\delta) + \eta \psi(x), \quad (3)$$

17, 18 y 19 de octubre de 2018

donde el término de fidelidad  $\phi(\mathcal{K}x, y^\delta): \mathcal{Y} \times \mathcal{Y} \rightarrow [0, \infty]$  “mide” la proximidad entre  $\mathcal{K}x$  y  $y^\delta$ , y  $\psi$  es un funcional no negativo definido en  $\mathcal{X}$  que codifica la información que se dispone a priori de la variable incógnita  $x$ . La solución  $x^\dagger$  se aproximará por un minimizante de  $\mathcal{J}_{\eta, y^\delta}(x)$ , el cual denotaremos con  $x_\eta^\delta$ . El parámetro  $\eta$  está asociado al compromiso entre el término de fidelidad  $\phi(\mathcal{K}x, y^\delta)$  y el término de regularización  $\psi(x)$ , y su selección constituye una de las mayores dificultades, siendo crucial para el éxito de un método de regularización.

Para la resolución de un problema inverso existe un enfoque diferente (y en principio totalmente disjunto) al determinístico hasta aquí presentado: el enfoque estocástico. Mientras que los métodos clásicos (determinísticos) producen una única estimación de la solución del problema, los métodos estadísticos dan como resultado una distribución de probabilidades que puede utilizarse para obtener estimaciones de la incógnita. Esta distribución describe el grado de conocimiento de la incógnita después de que la información contenida en los datos ha sido procesada e incorporada a tal distribución, y se denomina distribución de probabilidades *a-posteriori*.

En este enfoque, se considera que las variables involucradas en el modelo corresponden a variables aleatorias

absolutamente continuas y que, por ende, su distribución de probabilidades se representa por medio de una función de densidad. Entonces, desde este punto de vista, un problema inverso se expresa de la siguiente manera: *dada una observación  $y_{obs}$  de  $Y$  con  $p_y(y_{obs}) > 0$ , hallar la distribución de probabilidad condicional  $p(x|Y = y_{obs})$  de la variable  $X$ .*

Esta distribución puede obtenerse utilizando el Teorema de Bayes (Kaipio & Somersalo, 2005) del cual se sigue que

$$p_{post}(x) := p(x|y_{obs}) = \frac{p_{pr}(x) p(y_{obs}|x)}{p_y(y_{obs})} \\ \propto p_{pr}(x) p(y_{obs}|x).$$

Luego, la resolución de un problema inverso desde el enfoque estadístico puede resumirse en la realización de las siguientes tareas: **1)** basado en la información disponible de la variable incógnita  $X$ , hallar una distribución a-priori  $p_{pr}(x)$  que refleje dicha información; **2)** hallar la función de verosimilitud  $p(y_{obs}|x)$ ; **3)** desarrollar métodos para explorar la distribución *a-posteriori*. Es claro que la definición abstracta de la solución de un problema inverso como la distribución de probabilidad *a-posteriori* no resulta muy útil en la práctica y, en general, se utiliza el estimador *máximo a-posteriori*, definido como

$$x_{MAP} = \underset{x \in \mathbb{R}^m}{\operatorname{argmin}} p_{post}(x),$$

para aproximar la incógnita del problema.

Para motivar ahora la construcción de una regla de elección para el parámetro  $\eta$  en (3), discretizamos el problema (2) y obtenemos el modelo finito-dimensional

$$Kx = y^\delta, \quad (4)$$

con  $K \in \mathbb{R}^{n \times m}$ ,  $x \in \mathbb{R}^m$  e  $y^\delta \in \mathbb{R}^n$ , el cual será abordado desde el enfoque estadístico Bayesiano. Si consideramos ruido aditivo (es decir, el modelo para este caso resulta  $Kx + e = y^\delta$ ) y Gaussiano de media cero y matriz de covarianzas  $\sigma^2 I_n$ , obtenemos la función de verosimilitud

$$p(y^\delta|x, \tau) \propto \tau^{n/2} \exp\left(-\frac{\tau}{2} \|Kx - y^\delta\|_2^2\right),$$

con  $\tau = 1/\sigma^2$ .

Para codificar la información que se tiene de la variable de interés antes de recolectar los datos, consideraremos la densidad a-priori

$$p_{pr}(x|\lambda) \propto \lambda^{m'/2} \exp\left(-\frac{\lambda}{2} \|Lx\|_2^2\right),$$

donde la matriz  $L \in \mathbb{R}^{p \times m}$  está relacionada con la estructura de las interacciones entre los vecinos y típicamente corresponde a la discretización de un operador diferencial, el parámetro de escala  $\lambda$  está relacionado con la fuerza de dichas interacciones y  $m' = \operatorname{Rango}(L)$ . Desafortunadamente, los parámetros  $\lambda$  y  $\tau$  son desconocidos y es por esta razón que serán considerados variables aleatorias con distribución Gamma cuyas funciones de densidad (las que se conocen como hiper-priors ya que dan lugar a un hiper-modelo) son

$$p(\lambda) \propto \lambda^{\alpha_0 - 1} \exp(-\beta_0 \lambda)$$

y

$$p(\tau) \propto \tau^{\alpha_1-1} \exp(-\beta_1 \tau),$$

donde  $(\alpha_0, \beta_0)$  y  $(\alpha_1, \beta_1)$  son pares de parámetros conocidos y positivos.

Utilizando la regla de Bayes, la densidad *a-posteriori* para  $(X, \lambda, \tau)$  resulta

$$\begin{aligned} p(x, \lambda, \tau | y^\delta) &\propto p(y^\delta | x) p_{pr}(x | \lambda) p(\lambda) p(\tau) \\ &\propto \exp\left(-\frac{\tau}{2} \|Kx - y^\delta\|_2^2 - \frac{\lambda}{2} \|Lx\|_2^2 - \beta_0 \lambda - \beta_1 \tau\right) \\ &\quad + \left(\frac{m'}{2} + \alpha_0 - 1\right) \ln \lambda \\ &\quad + \left(\frac{n}{2} + \alpha_1 - 1\right) \ln \tau. \end{aligned}$$

En este caso obtener el estimador MAP es equivalente a minimizar el *funcional de Tikhonov-Phillips generalizado aumentado*

$$J(x, \lambda, \tau) = \tau \phi(\mathcal{K}x, y^\delta) + \lambda \psi(x) + \beta'_0 \lambda + \beta'_1 \tau - \alpha'_0 \ln \lambda - \alpha'_1 \ln \tau, \quad (5)$$

definido para  $(x, \lambda, \tau) \in \mathcal{X} \times \mathbb{R}^+ \times \mathbb{R}^+$ , el cual extiende el funcional dado en (3). Para simplificar la notación a partir de aquí denotaremos con  $\beta_0$  y  $\beta_1$  a  $\beta'_0$  y  $\beta'_1$ , respectivamente.

Ito et al. propusieron una regla de elección de parámetro a partir del concepto de punto crítico que se define a continuación.

**Definición:** Se dice que un elemento  $(x^*, \lambda^*, \tau^*) \in \mathcal{X} \times \mathbb{R}^+ \times \mathbb{R}^+$  es un *punto crítico* del funcional (5) si satisface el siguiente sistema de optimalidad generalizado

$$\begin{cases} x^* = \operatorname{argmin}_{x \in \mathcal{X}} \{ \phi(\mathcal{K}x, y^\delta) + \lambda^* (\tau^*)^{-1} \psi(x) \} \\ \psi(x^*) + \beta_0 - \alpha'_0 \frac{1}{\lambda^*} = 0 \\ \phi(\mathcal{K}x^*, y^\delta) + \beta_1 - \alpha'_1 \frac{1}{\tau^*} = 0. \end{cases}$$

Es oportuno observar que esta definición de punto crítico coincide con la que se realiza habitualmente anulando las derivadas parciales

$$\begin{aligned} \frac{\partial}{\partial \lambda} J(x^*, \lambda^*, \tau^*) &= \psi(x^*) + \beta_0 - \alpha'_0 \frac{1}{\lambda^*} \quad y \\ \frac{\partial}{\partial \tau} J(x^*, \lambda^*, \tau^*) &= \phi(\mathcal{K}x^*, y^\delta) + \beta_1 - \alpha'_1 \frac{1}{\tau^*}. \end{aligned}$$

Luego, resulta

$$\lambda^* = \frac{\alpha'_0}{\psi(x^*) + \beta_0},$$

y

$$\tau^* = \frac{\alpha'_1}{\phi(\mathcal{K}x^*, y^\delta) + \beta_1}.$$

Por lo tanto, de la definición de punto crítico se sigue que el parámetro  $\eta^*$  está determinado por

$$\eta^* = \lambda^* \tau^{*-1} = \frac{\alpha'_0}{\psi(x^*) + \beta_0} \cdot \frac{\phi(\mathcal{K}x^*, y^\delta) + \beta_1}{\alpha'_1}. \quad (6)$$

De la teoría clásica (determinística) de regularización se sabe que una regla de elección de parámetro *a-priori*  $\eta(\delta)$  debe verificar que  $\lim_{\delta \rightarrow 0} \eta(\delta) = 0$  y  $\lim_{\delta \rightarrow 0} \frac{\delta^2}{\eta(\delta)} = 0$ , para asegurar que forma parte de un método de regularización, es decir, para asegurar que la solución aproximada converja a la solución exacta cuando el nivel de ruido tienda a cero. Debido a que la regla (6) no verifica el último límite, se modifica de la siguiente manera

$$\begin{aligned} \eta^* = \lambda^* \tau^{*-1} &= \frac{\alpha'_0}{\psi(x^*) + \beta_0} \\ &\quad \cdot \left( \frac{\phi(\mathcal{K}x^*, y^\delta)}{\alpha'_1} \right)^{1-d} \\ &= \alpha \frac{\phi(\mathcal{K}x^*, y^\delta)^{1-d}}{\psi(x^*) + \beta_0}, \quad (7) \end{aligned}$$

donde  $0 < d < 1$ ,  $\alpha = \frac{\alpha'_0}{(\alpha'_1)^{1-d}}$  y la constante  $\beta_1$  se elimina por considerarse cercana a cero.

Para implementar numéricamente la regla (7) en ejemplos de procesamiento de señales e imágenes, se propone un algoritmo que consta de los siguientes pasos:

Elegir una posición inicial  $\eta_0 > 0$  y fijar  $k = 0$ .

**I.** Hallar el minimizante  $x_{k+1}$  del funcional de Tikhonov-Phillips:

$$x_{k+1} = \operatorname{argmin}_{x \in \mathcal{X}} \{ \phi(\mathcal{K}x, y^\delta) + \eta_k \psi(x) \};$$

**II.** Actualizar el parámetro  $\eta_{k+1}$ :

$$\eta_{k+1} = \alpha \frac{\phi(\mathcal{K}x_{k+1}, y^\delta)^{1-d}}{\psi(x_{k+1}) + \beta_0};$$

**III.** Chequear el criterio de parada establecido. Si no se alcanza, fijar  $k = k + 1$  y repetir los pasos **I** y **II**.

Para esta regla de elección de parámetro, en Ito et al., se presenta un resultado de cota para el error a-posteriori considerando que  $\mathcal{Y}$  es un espacio de Hilbert,  $\phi(\mathcal{K}x, y^\delta) = \frac{1}{2} \|\mathcal{K}x - y^\delta\|^2$  y  $\psi$  un funcional convexo. En este trabajo, estamos interesados en obtener una cota para dicho error pero generalizando el término de fidelidad de manera tal que  $\phi(\mathcal{K}x, y^\delta) = \frac{1}{p} \|\mathcal{K}x - y^\delta\|^p$  con  $p > 1$ , y bajo las mismas hipótesis sobre  $\mathcal{Y}$  y  $\psi$ . De esta manera, el funcional resulta

$$\mathcal{J}_{\eta, y^\delta}^p(x) = \frac{1}{p} \|\mathcal{K}x - y^\delta\|^p + \eta \psi(x),$$

y denotaremos con  $x_\eta^\delta$  a su minimizante.

Análogamente a Ito et al., para medir el error a-posteriori utilizaremos la distancia de Bregman entre  $x^\dagger$  y  $x_\eta^\delta$ . Para introducir la definición de distancia de Bregman se requiere el concepto de subdiferencial del operador  $\psi$ .

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

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

$$d_{\xi^\dagger}^\psi(x^\dagger, x) = \psi(x) - \psi(x^\dagger) - [\xi^\dagger, x - x^\dagger],$$

para todo  $x \in \mathcal{X}$ .

Es fácil ver que si  $\mathcal{X}$  es un espacio de Hilbert y  $\psi(x) = \|x\|^2$ , entonces resulta que la distancia de Bregman entre  $x$  e  $y$  coincide con  $\|x - y\|^2$ .

Finalmente, es necesario introducir la definición de solución  $\psi$ -minimizante de un problema inverso.

**Definición:** Decimos que  $\tilde{x} \in \mathcal{X}$  es una *solución  $\psi$ -minimizante* de (1) si  $\mathcal{K}\tilde{x} = y$  y  $\psi(\tilde{x}) \leq \psi(x)$ , para todo  $x \in \mathcal{X}$  tal que  $\mathcal{K}x = y$ .

El siguiente Teorema presenta una cota para la distancia de Bregman entre  $x^\dagger$  y  $x_\eta^\delta$ .

**Teorema (Cota de error a-posteriori):** Sea  $x^\dagger$  una solución  $\psi$ -minimizante de (1) y supongamos que se satisface la siguiente condición fuente: existe  $w \in \mathcal{Y}$  tal que  $\xi = \mathcal{K}^*w \in \partial\psi(x^\dagger)$ . Para cada minimizante  $x_\eta^\delta$  del funcional  $J_{\eta,y^\delta}^p$ , si  $\|y - y^\delta\| \leq \delta$  se tiene que

- Si  $p > 1$  y  $q$  es tal que  $\frac{1}{q} + \frac{1}{p} = 1$  entonces

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

- Si  $p = 1$  y  $\eta\|w\| < 1$  entonces

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

**Demostración:**

Como  $x_\eta^\delta$  es un minimizante del funcional  $J_{\eta,y^\delta}^p$  se tiene que

$$\begin{aligned} & \frac{1}{p}\|\mathcal{K}x_\eta^\delta - y^\delta\|^p + \eta\psi(x_\eta^\delta) \\ & \leq \frac{1}{p}\|\mathcal{K}x^\dagger - y^\delta\|^p + \eta\psi(x^\dagger) \\ & = \frac{1}{p}\delta^p + \eta\psi(x^\dagger). \end{aligned}$$

Luego, de la definición de distancia de Bregman, la condición fuente, la desigualdad de Cauchy-Schwartz y del hecho que  $\|y - y^\delta\| \leq \delta$ , resulta

$$\begin{aligned} & \frac{1}{p}\|\mathcal{K}x_\eta^\delta - y^\delta\|^p + \eta d_{\xi^\dagger}^\psi(x_\eta^\delta, x^\dagger) \\ & \leq \frac{1}{p}\delta^p - \eta[\xi, x_\eta^\delta - x^\dagger] \end{aligned}$$

$$\begin{aligned} & = \frac{1}{p}\delta^p + \eta\langle w, \mathcal{K}x^\dagger - \mathcal{K}x_\eta^\delta \rangle \\ & \leq \frac{1}{p}\delta^p + \eta\|w\|\|\mathcal{K}x^\dagger - \mathcal{K}x_\eta^\delta\| \\ & \leq \frac{1}{p}\delta^p + \eta\|w\|\delta + \eta\|w\|\|\mathcal{K}x_\eta^\delta - y^\delta\|. \end{aligned}$$

Finalmente,

$$\begin{aligned} & \frac{1}{p}\|\mathcal{K}x_\eta^\delta - y^\delta\|^p - \eta\|w\|\|\mathcal{K}x_\eta^\delta - y^\delta\| \\ & \quad + \eta d_{\xi^\dagger}^\psi(x_\eta^\delta, x^\dagger) \\ & \leq \frac{1}{p}\delta^p + \eta\|w\|\delta. \quad (8) \end{aligned}$$

- Si  $p > 1$ : por la desigualdad de Young se tiene

$$\begin{aligned} & -\eta\|w\|\|\mathcal{K}x_\eta^\delta - y^\delta\| \\ & \geq -\frac{1}{p}\|\mathcal{K}x_\eta^\delta - y^\delta\|^p - \frac{1}{q}(\eta\|w\|)^q. \quad (9) \end{aligned}$$

Reemplazando (9) en (8) resulta

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

y, de esta manera, se tiene

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

- Si  $p = 1$ : de (8) resulta que

$$\begin{aligned} & (1 - \eta\|w\|)\|\mathcal{K}x_\eta^\delta - y^\delta\| + \eta d_{\xi^\dagger}^\psi(x_\eta^\delta, x^\dagger) \\ & \leq \delta + \eta\|w\|\delta, \end{aligned}$$

y como  $\eta\|w\| < 1$  se obtiene

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

En Ito et al. se prueba, para el caso  $p = 2$ , la siguiente desigualdad

$$d_{\xi^\dagger}^\psi(x_\eta^\delta, x^\dagger) \leq \frac{\delta^2}{\eta} + \eta\|w\|^2. \quad (10)$$

Es importante observar ahora que para  $p = 2$ , del Teorema anterior se sigue inmediatamente que

$$d_{\xi}^{\psi}(x_{\eta}^{\delta}, x^{\dagger}) \leq \frac{\frac{1}{2}\delta^2 + \eta\|w\|\delta + \frac{1}{2}(\eta\|w\|)^2}{\eta}$$

$$= \frac{1}{2}\left(\frac{\delta}{\sqrt{\eta}} + \sqrt{\eta}\|w\|\right)^2. \quad (11)$$

Es bien sabido que  $\frac{1}{2}(a+b)^2 \leq a^2 + b^2$ , para todo  $a, b \geq 0$ , y como consecuencia, se tiene que la cota obtenida en (11) es menor o igual a la cota dada en (10).

Es oportuno mencionar que la cota en (11) coincide con la que M. Burger y S. Osher obtuvieron en el 2004 para el término de fidelidad cuadrático  $\frac{1}{2}\|\mathcal{K}x - y^{\delta}\|_2^2$ .

El siguiente corolario presenta una cota para el error a-posteriori para la regla de elección de parámetro (7).

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

$\|\mathcal{K}x_{\eta^*}^{\delta} - y^{\delta}\|$ , entonces

- Si  $p > 1$ :

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

- Si  $p = 1$ :

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

**Demostración:** El resultado se sigue inmediatamente reemplazando  $\eta^*$  en la cota obtenida en el Teorema anterior. ■

A continuación, presentamos diferentes ejemplos numéricos en donde se

implementó este algoritmo en procesamiento de señales e imágenes, utilizando como término de fidelidad  $\phi(\mathcal{K}x, y^{\delta}) = \frac{1}{2}\|\mathcal{K}x - y^{\delta}\|_2^2$ .

### Caso unidimensional: Aplicaciones a deconvolución de señales

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

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

donde  $k$  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  $y = Kx$ , donde  $x, y \in \mathbb{R}^n$  y la matriz  $K \in \mathbb{R}^{n \times n}$  resulta de la discretización del núcleo  $k$ . En los ejemplos presentados, se consideró  $n = 100$  y el dato fue contaminado mediante un ruido Gaussiano de la siguiente manera

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

donde  $\varepsilon = 0.01$  y  $\xi_i$  tiene una distribución Gaussiana estándar, de lo cual resulta que  $\varepsilon$  y la varianza  $\sigma^2$  del error están relacionadas través de la fórmula  $\sigma = \varepsilon \max_i\{|y_i|\}$ .



Ejemplo 1:  $\psi(x) = \|x\|_1$

Este ejemplo se basa en el presentado en Ito et al. (pp. 1539-1542). Aquí, utilizamos el núcleo de convolución  $k(s, t) = \varphi(s - t)$ , con  $\varphi(t) = \left(1 + \cos \frac{\pi}{3} t\right) \chi_{|t-s| < 3}$  y  $\chi$  denota la función característica. La señal original está dada por

$$x(t) = -1.5 \chi_{[-5.12, 5]} + \chi_{[-0.92, 0.8]} - 1.2 \chi_{[0.9, 1.02]} - \chi_{[4.1, 4.22]},$$

y el intervalo de integración correspondiente es  $I = [-6, 6]$ .

En la Figura 1 se muestra el dato contaminado con ruido Gaussiano mediante el procedimiento descrito previamente.

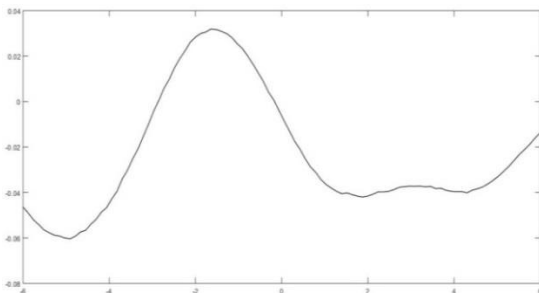


Figura 1: Dato contaminado con ruido Gaussiano.

Utilizando el algoritmo propuesto con parámetros  $\eta_0 = 10^{-5}$ ,  $\alpha = 0.1$ ,  $d = \frac{1}{3}$ ,  $\beta_0 = 10^{-4}$ , luego de 8 iteraciones se obtuvo el parámetro de regularización óptimo  $\eta^* = 0.694 \times 10^{-4}$ . En la Figura 2 se presentan la señal original  $x$  (—) y la señal  $x_{\eta^*}^\delta$  (---) obtenida mediante el algoritmo. Se puede observar que el procedimiento logra

recuperar las características de la señal original.

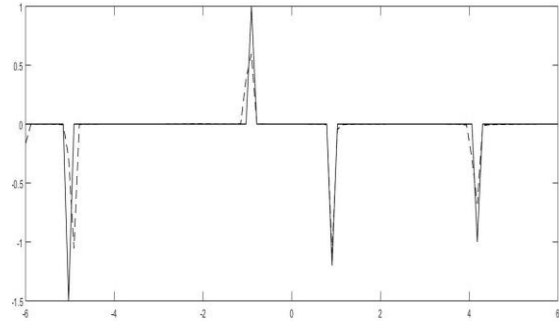


Figura 2: Señal original (—) y solución (---) obtenida mediante el algoritmo con  $\eta^* = 0.694 \times 10^{-4}$ .

Ejemplo 2:  $\psi(x) = \frac{1}{2} \|x\|_2^2$

Este ejemplo se basa en el presentado en Ito et al. Aquí, utilizamos el núcleo de convolución  $k(s, t) = (\cos s + \cos t) \left(\frac{\sin u}{u}\right)^2$ , con  $u(s, t) = \pi(\sin s + \sin t)$  y la señal original dada por

$$x(t) = 2e^{-6\left(t-\frac{4}{5}\right)^2} + e^{-2\left(t+\frac{1}{2}\right)^2},$$

definida en el intervalo  $I = \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ .

En la Figura 3 se muestra el dato contaminado con ruido Gaussiano con el procedimiento descrito previamente.

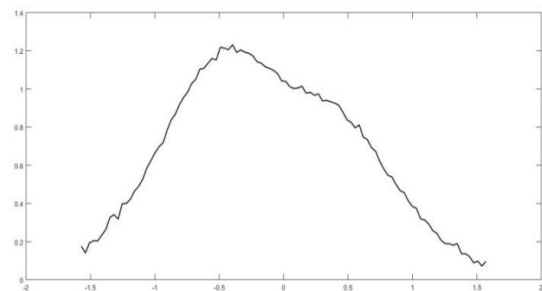


Figura 3: Dato contaminado con ruido Gaussiano.

Utilizando el algoritmo propuesto con parámetros  $\eta_0 = 10^{-5}$ ,  $\alpha = 1$ ,  $d = \frac{1}{3}$ ,  $\beta_0 = 10^{-4}$ , luego de 9 iteraciones se obtuvo el parámetro de regularización óptimo  $\eta^* = 0.753 \times 10^{-3}$ . En la Figura 4 se presentan la señal original  $x$  (—) y la señal  $x_{\eta^*}^{\delta}$  (---) obtenida mediante el algoritmo, donde puede observarse que el procedimiento logra una buena aproximación de la señal original.

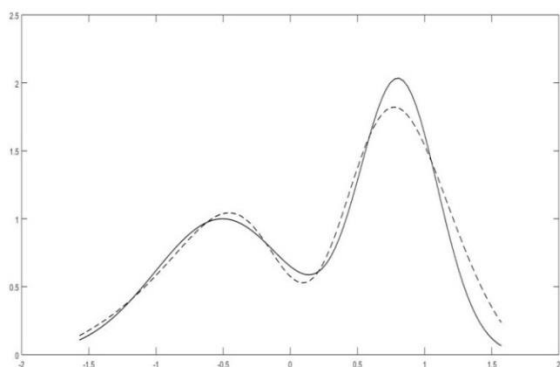


Figura 4: Señal original (—) y solución (---) obtenida mediante el algoritmo con  $\eta^* = 0.753 \times 10^{-3}$ .

#### Caso bidimensional: Aplicaciones a restauración de imágenes

Análogamente al caso unidimensional, un modelo matemático muy general para el proceso de degradación de imágenes está dado por la siguiente ecuación integral

$$g(s, t) = \frac{1}{2\pi\sigma_h\sigma_v} \iint_I \exp\left(-\frac{1}{2}\left(\frac{s-s'}{\sigma_h}\right)^2 - \frac{1}{2}\left(\frac{t-t'}{\sigma_v}\right)^2\right) f(s', t') ds' dt',$$

donde  $f$  y  $g$  son las imágenes original y degradada, respectivamente, las desviaciones estándar  $\sigma_h$  y  $\sigma_v$  caracterizan la degradación en la dirección horizontal y vertical, respectivamente, y la región de integración está dada por  $[0,1] \times [0,1]$ . El modelo discretizado con ruido aditivo está dado por  $y = Kx + e$ , donde, análogamente al caso unidimensional, la matriz  $K$  de orden  $N \times N$  está asociada a la discretización del núcleo Gaussiano.  $x, y, e \in \mathbb{R}^N$  con  $N = n^2$ , debido a que la imagen original es de tamaño  $n \times n$ .

#### Ejemplo 3: $\psi(x) = \|x\|_1$

El siguiente ejemplo se basa en el presentado en Carrió (2017) donde se considera una imagen original de  $60 \times 60$  píxeles con cuatro fuentes focales. La matriz del modelo fue obtenida utilizando un núcleo Gaussiano con  $\sigma_h = \sigma_v = 2$  y condición de borde reflexiva (Hansen, Nagy & Leary, 2006). A su vez, se consideró el ruido  $E$  con distribución  $N(0, \sigma_e I_N)$  con  $\sigma_e = 0.01 \|Y\|_{\infty}$ . En la Figura 5 se muestran las imágenes original y borrosa con ruido.

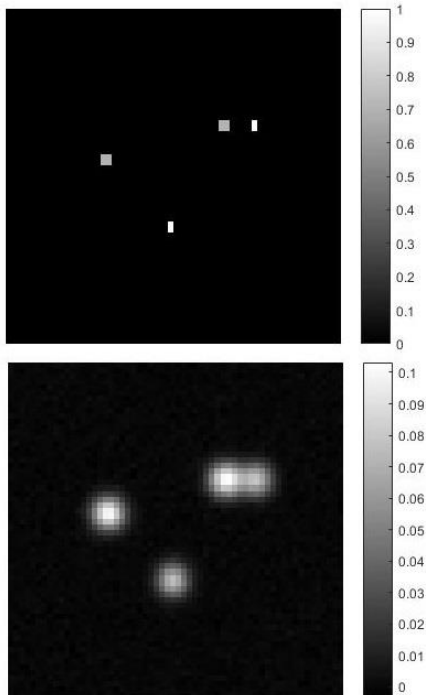


Figura 5: Imagen original (arriba) y borrosa con ruido (abajo).

Utilizando el algoritmo propuesto con parámetros  $\eta_0 = 10^{-5}$ ,  $\alpha = 0.05$ ,  $d = \frac{1}{3}$ ,  $\beta_0 = 10^{-10}$ , luego de 6 iteraciones se obtuvo el parámetro de regularización óptimo  $\eta^* = 1.2135 \times 10^{-4}$  y la solución que se muestra debajo.

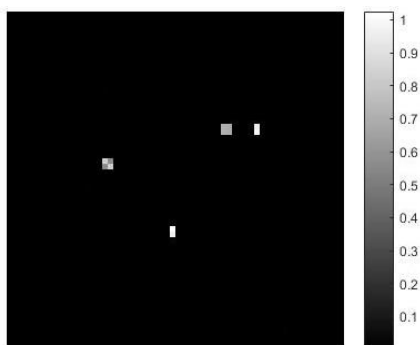


Figura 6: Solución obtenida mediante el algoritmo con  $\eta^* = 1.2135 \times 10^{-4}$ .

Se puede observar en la Figura 6 que el procedimiento logra detectar la ubicación y la forma de los objetos focales presentes en la imagen original.

### Conclusiones

Se estudió una regla de elección de parámetro de regularización para un funcional de Tikhonov-Phillips generalizado. Para dicha regla, se obtuvo una cota para el error a-posteriori entre la solución exacta del problema y la solución regularizada obtenida, considerando como término de fidelidad  $\frac{1}{p} \|Kx - y\|^p$  y como penalizante un funcional convexo  $\psi$ . De esta forma, no sólo se logró generalizar el resultado en Ito et al. sino que también se obtuvo una cota menor a la cota dada por estos autores. Finalmente, se implementó un algoritmo numérico para la elección del parámetro de regularización, el cual fue aplicado a ejemplos de deconvolución de señales e imágenes. En dichos ejemplos, la solución obtenida a partir del parámetro óptimo logró recuperar las características de la solución exacta del problema.

### Bibliografía

- Burger, M. & Osher, S. (2004). Convergence rates of convex variational regularization. *Inverse Problems*. (20). 1411–1421.
- Carrió, M. J. (2017, octubre). *Un enfoque estadístico para la detección de objetos focales en un problema*

Universidad Nacional de Cuyo | Mendoza | Argentina  
*inverso de restauración de imágenes*. XXI Encuentro de Jóvenes Investigadores de la Universidad Nacional del Litoral, Santa Fe. Recuperado de <http://bibliotecavirtual.unl.edu.ar:8080/colecciones/handle/123456789/8823>.

Engl, H. W, Hanke, M., & Neubauer, A. (1996). Regularization of inverse problems. *Mathematics and its Applications*. (375).

Hansen, P.C., Nagy, J.G., & O'Leary, D.P. (2006). *Deblurring Images - Matrices, Spectra and Filtering*. SIAM: Philadelphia.

Ito, K., Jin, B. & Zou, J. (2011). A new choice rule for regularization parameters in Tikhonov regularization. *Applicable Analysis*. (90). 1521–1544.

Kaipio, J. & Somersalo, E. (2005). Statistical and computational inverse problems. *Applied Mathematical Sciences*. (160).

Lagendijk, R. L., & Biemond, J. (1991). *Iterative Identification and Restoration of Images*. Kluwer, Boston, Dordrecht, London.

Phillips, D. L. (1962). A technique for the numerical solution of certain integral equations of the first kind. *J. Assoc. Comput.*(9). 84–97.

17, 18 y 19 de octubre de 2018

Tikhonov, A. N. (1963a). Regularization of incorrectly posed problems. *Soviet Math.* (4). 1624–1627.

Tikhonov, A. N. (1963b) Solution of incorrectly formulated problems and the regularization method. *Soviet Math.* (4). 1035–1038.

### Financiamiento

Este trabajo ha sido subvencionado por la Universidad Nacional del Litoral, a través del proyecto CAI+D-UNL 2016, PIC 50420150100036LI.