

*Facultad
de
Ciencias*

**Estudio de técnicas de “unfolding” en
física experimental
(Unfolding techniques in experimental
physics)**

Trabajo de Fin de Grado
para acceder al

GRADO EN MATEMÁTICAS

Autor: Máximo García Lomas

Director: Francisco Matorras Weinig

Septiembre - 2024

Resumen Este Trabajo de Fin de Grado aborda el problema del *unfolding* en estadística, centrado en la estimación de distribuciones verdaderas a partir de observaciones distorsionadas por dispositivos de medición imperfectos. A lo largo del estudio, se exploraron y compararon diversas técnicas, como los mínimos cuadrados, la regularización de Tikhonov y el método iterativo con *early stopping*. Cada metodología presenta ventajas y limitaciones que dependen de la condición del problema y de la calidad de los datos. Se destacó la importancia de una adecuada selección de los coeficientes de regularización y del número de iteraciones para balancear el sesgo y la varianza, y evitar tanto el sobreajuste como la pérdida de información relevante. Los resultados obtenidos subrayan la necesidad de un enfoque adaptativo, que considere las características específicas de cada problema para garantizar estimaciones precisas y estables.

Abstract This Bachelor's Thesis addresses the problem of *unfolding* in statistics, focusing on estimating true distributions from observations distorted by imperfect measurement devices. Throughout the study, various techniques were explored and compared, including least squares, Tikhonov regularization, and the iterative method with *early stopping*. Each methodology presents advantages and limitations depending on the problem's condition and the quality of the data. The importance of proper selection of regularization coefficients and the number of iterations was highlighted to balance bias and variance, avoiding both overfitting and loss of relevant information. The results obtained underscore the need for an adaptive approach that considers the specific characteristics of each problem to ensure accurate and stable estimates.

Palabras Clave Unfolding estadístico, Regularización de Tikhonov, Iteraciones de D'Agostini, Sesgo, Varianza.

Índice general

1. Introducción	3
2. Marco Teórico	5
2.1. Formulación del problema	5
2.2. Ajuste por máxima verosimilitud	6
2.3. Problema de mínimos cuadrados	7
2.4. Regularización de Tikhonov	8
2.5. Método iterativo con 'early stopping'.	9
2.6. Cálculo del sesgo, la varianza y la correlación	10
2.7. Problemas y dificultades del <i>Unfolding</i>	13
3. Presentación del problema	14
3.1. Elección de K y Justificación	14
3.2. Ajuste por máxima verosimilitud	15
4. Método de Tikhonov	17
4.1. Curvatura	17
4.2. SVD unfolding	18
4.3. TUnfold	20
4.4. Elección del coeficiente de regularización δ	21
5. Método iterativo con 'early stopping'	27
5.1. Una iteración	27
5.2. Más iteraciones	28
6. Conclusiones	34
A. Anexo A: Regularización de Tikhonov. Gráficas adicionales.	36
B. Anexo B: Método iterativo con 'early stopping'. Gráficas adicionales.	39
C. Anexo C: Intercambio sesgo-varianza (<i>Bias-Variance Trade Off</i>). Gráficas adicionales.	44

Capítulo 1

Introducción

El *unfolding* en estadística se refiere al problema de estimar la distribución a nivel de partícula de alguna cantidad física de interés a partir de observaciones distorsionadas por un dispositivo de medición imperfecto. Se trabaja con espectros, con espectro nos referimos a una representación de la distribución de eventos físicos observados o medidos, generalmente en función de una variable física como la energía, el momento, o alguna otra magnitud relevante. En mediciones experimentales, como las realizadas en física de partículas, el espectro observado suele estar afectado por la resolución finita de los detectores, lo que significa que el espectro verdadero se ve “desenfocado” o “difuminado”.

El problema de *unfolding* es precisamente inferir el espectro verdadero (o el “espectro” que se mediría sin los efectos de resolución del detector) a partir del espectro observado que ha sido distorsionado por estas limitaciones. Este proceso implica resolver un problema inverso, donde se intenta corregir las distorsiones y recuperar el espectro verdadero a partir de las observaciones.

Generalmente es usado para lograr uno o más de los siguientes objetivos:

1. Comparación de la medición con teorías futuras.
2. Comparación de experimentos con diferentes respuestas.
3. Entrada a un análisis posterior.
4. Análisis exploratorio de datos.
5. Problemas de inversión, aparentemente sencillos, pero que en casos realistas son mal definidos o planteados.

La aproximación a los problemas de *unfolding* suele ser mediante simulación. Sin embargo, vamos a abordarlo de otra forma, más analítica, que aunque implica aproximaciones, permite entender mejor los resultados. Además, en muchos casos el problema resulta ser lineal por lo que la solución es exacta.

Este trabajo se organiza de la siguiente manera:

En el **Capítulo 2**, se presenta el marco teórico esencial para entender el problema del *unfolding* en estadística. Se comienza con la formulación del problema, describiendo la relación matemática entre el espectro verdadero y el espectro distorsionado debido a las imperfecciones en la medición. A continuación, se exploran

diversas metodologías para abordar el *unfolding*, incluyendo el ajuste por máxima verosimilitud, el uso de mínimos cuadrados, y la regularización de Tikhonov. Se introduce también el método iterativo con *early stopping*, explicando su aplicación en situaciones donde el problema es mal condicionado. Finalmente, se discuten los cálculos necesarios para evaluar el sesgo, la varianza y la correlación en las estimaciones obtenidas.

En el **Capítulo 3**, se aplica la teoría a un caso práctico, comenzando con la elección y justificación de la matriz de dispersión K , un componente crucial para modelar correctamente el problema. Este capítulo detalla cómo se pueden aplicar los métodos de mínimos cuadrados sin regularización para resolver el problema de *unfolding*. Se examinan las consecuencias de utilizar diferentes configuraciones de la matriz K y cómo estas afectan a la precisión y estabilidad de las estimaciones obtenidas.

En el **Capítulo 4**, se exploran en detalle las metodologías de regularización, con especial énfasis en la regularización de Tikhonov y el uso de descomposiciones en valores singulares (*SVD unfolding*). Se analiza cómo estas técnicas pueden mejorar la estabilidad de las soluciones en problemas mal condicionados, y se discute la importancia de una adecuada selección de δ para minimizar el sesgo y la varianza. Este capítulo también incluye un análisis comparativo de diferentes enfoques de regularización, destacando sus ventajas y limitaciones en distintos escenarios experimentales.

El **Capítulo 5** se centra en el método iterativo con 'early stopping', una técnica que permite evitar el sobreajuste en problemas mal condicionados. Se compara su eficacia con la regularización de Tikhonov y se analizan los resultados obtenidos a través de una o varias iteraciones. En particular, se enfatiza que en situaciones donde el problema está bien condicionado, el estimador de máxima verosimilitud es adecuado sin necesidad de *early stopping*, mientras que en problemas mal condicionados, detener las iteraciones en el momento oportuno es crucial para obtener estimaciones fiables.

Finalmente, en el **Capítulo 6**, se presentan las conclusiones del trabajo, destacando las circunstancias en las que es recomendable aplicar técnicas de regularización para resolver problemas de *unfolding* mal condicionados. Se subraya la importancia de ajustar cuidadosamente los parámetros de regularización y el número de iteraciones para obtener resultados robustos, evitando tanto el sobreajuste como la subestimación de la incertidumbre en los datos.

Con esta estructura, el trabajo busca proporcionar una comprensión integral del problema del *unfolding*, combinando un enfoque teórico riguroso con aplicaciones prácticas detalladas para la resolución efectiva de problemas estadísticos complejos.

Capítulo 2

Marco Teórico

Comprender las herramientas matemáticas y estadísticas que subyacen al proceso de *unfolding* es esencial para abordar los desafíos que este presenta. En el corazón de esta problemática se encuentran conceptos como la verosimilitud, los mínimos cuadrados y la regularización, cada uno jugando un papel crucial en la corrección de las distorsiones introducidas por los dispositivos de medición.

En este capítulo, exploraremos el marco teórico que sustenta el *unfolding*, desglosando las ecuaciones y los métodos que permiten transformar un espectro alterado en una estimación confiable del espectro verdadero. La mayoría de los resultados discutidos en el apartado 2.1 están basados en el trabajo *Uncertainty Quantification in Unfolding Elementary Particle Spectra at the Large Hadron Collider* [1], mientras que los del apartado 2.2 y 2.3 se fundamentan en *Unfolding: A Statistician's Perspective* [2]. Por otro lado, los conceptos y resultados presentados en el apartado 2.4 están basados en el trabajo clásico de A. N. Tikhonov, *Solution of Incorrectly Formulated Problems and the Regularization Method* [3](1963). Desde la formulación del problema hasta las estrategias más avanzadas, como la regularización de Tikhonov y los métodos iterativos, este capítulo sienta las bases matemáticas y conceptuales necesarias para avanzar en el análisis del *unfolding*.

2.1. Formulación del problema

Sea f el espectro verdadero y g el espectro alterado a nivel del detector. Denotemos el espacio verdadero por T y el espacio alterado por S (ambos tomados como intervalos en la recta real para simplificar). f y g son las funciones de intensidad del espectro. Los dos espectros están relacionados por

$$g(s) = \int_T k(s, t) f(t) dt,$$

donde el núcleo alterado k representa la respuesta del detector o la convolución y está dado por

$$k(s, t) = p(Y = s \mid X = t, X \text{ observado})P(X \text{ observado} \mid X = t),$$

donde X es un evento verdadero e Y el evento alterado correspondiente.

El problema generalmente se discretiza utilizando histogramas, se discretiza la variable y cuenta el número de observaciones en los rangos correspondientes (a veces también se utilizan splines). Sean $\{T_i\}_{i=1}^p$ y $\{S_i\}_{i=1}^n$ particiones del espacio verdadero T y del espacio suavizado S . El histograma alterado $y = [y_1, \dots, y_n]^T$ tiene una media

$$\mu = \left[\int_{S_1} g(s) ds, \dots, \int_{S_n} g(s) ds \right]^T$$

Y la cantidad de interés es:

$$\lambda = \left[\int_{T_1} f(t) dt, \dots, \int_{T_p} f(t) dt \right]^T$$

Los histogramas con las medias están relacionados por $\mu = K\lambda$, donde los elementos de la matriz de respuesta K están dados por

$$K_{i,j} = \frac{\int_{S_i} \int_{T_j} k(s,t) f(t) dt ds}{\int_{T_j} f(t) dt} = P(\text{evento alterado en el bin } i \mid \text{evento verdadero en el bin } j)$$

El modelo estadístico discretizado se convierte en

$$y \sim \text{Poisson}(K\lambda)$$

Y lo que deseamos es hacer inferencias sobre λ bajo este modelo.

La primera idea que surgen observándolo es hacer un ajuste por máxima verosimilitud para obtener el estimador λ .

2.2. Ajuste por máxima verosimilitud

La función de verosimilitud en el unfolding es:

$$L(\lambda) = p(y \mid \lambda) = \prod_{i=1}^n p(y_i \mid \lambda) = \prod_{i=1}^n \frac{\left(\sum_{j=1}^p K_{ij} \lambda_j \right)^{y_i}}{y_i!} e^{-\sum_{j=1}^p K_{ij} \lambda_j}, \quad \lambda \in \mathbb{R}_+^n$$

Esta función utiliza nuestro modelo de regresión de Poisson para vincular las observaciones y con el λ desconocido. La función de verosimilitud juega un papel clave en todos los métodos razonables de unfolding. En la mayoría de los problemas estadísticos, el máximo de la función de verosimilitud, o equivalentemente el máximo de la función de verosimilitud logarítmica, proporciona una buena estimación del valor desconocido λ .

En problemas mal condicionados, esto generalmente no es el caso, pero la solución de máxima verosimilitud aún proporciona un buen punto de partida.

Cualquier histograma que maximice la verosimilitud logarítmica del problema de unfolding se llama estimador de máxima verosimilitud $\hat{\lambda}_{\text{MLE}}$ de λ . Por lo tanto, queremos resolver:

$$\max_{\lambda \in \mathbb{R}_+^n} \log p(y \mid \lambda) = \sum_{i=1}^n \left[y_i \log \left(\sum_{j=1}^p K_{ij} \lambda_j \right) - \sum_{j=1}^p K_{ij} \lambda_j \right] + \text{const}$$

Teorema 2.2.1 (Vardi et al. (1985)) [4] *Supongamos que $K_{ij} > 0$ y $y \neq 0$. Entonces, se cumple lo siguiente para la verosimilitud logarítmica $\log p(y | \lambda)$ del problema de unfolding:*

1. *La verosimilitud logarítmica tiene un máximo.*
2. *La verosimilitud logarítmica es cóncava y, por lo tanto, todos los máximos son máximos globales.*
3. *El máximo es único si y solo si las columnas de K son linealmente independientes.*

Entonces, si las columnas de K son linealmente independientes, existe un único estimador de máxima verosimilitud logarítmica. Queda ver como obtenerlo.

Proposición 2.2.2 *Sea K una matriz cuadrada invertible y supongamos que $\hat{\lambda} = K^{-1}y \geq 0$. Entonces, $\hat{\lambda}$ es el estimador de máxima verosimilitud (MLE) de λ .*

Es decir, la inversión de la matriz nos proporciona el MLE si K es invertible y la estimación resultante es positiva.

Cabe señalar que este resultado es más restrictivo de lo que puede parecer: K a menudo no es cuadrada; incluso si K fuera cuadrada, a menudo no es invertible; y, aunque K fuera invertible, $K^{-1}y$ a menudo contiene valores negativos.

2.3. Problema de mínimos cuadrados

Nosotros consideraremos $y \sim N(K\lambda, 1)$ como se explicará en el **Apartado 3.1**, con esta aproximación, el problema anterior se convierte en el siguiente problema de mínimos cuadrados:

$$\min_{\lambda \in \mathbb{R}^p} \|K\lambda - y\|^2,$$

donde $K \in \mathbb{R}^{n \times p}$, $\lambda \in \mathbb{R}^p$ y $y \in \mathbb{R}^n$.

Este problema siempre tiene una solución, aunque puede no ser única. Una solución siempre está dada por la pseudoinversa de Moore–Penrose de K :

$$\hat{\lambda}_{LS} = K^\dagger y.$$

Cuando hay múltiples soluciones, la pseudoinversa da aquella con la norma más pequeña. Cuando K tiene rango completo de columnas, la solución es única. En este caso, la pseudoinversa está dada por:

$$K^\dagger = (K^\top K)^{-1} K^\top.$$

Por lo tanto, la solución de mínimos cuadrados es:

$$\hat{\lambda}_{LS} = (K^\top K)^{-1} K^\top y.$$

Pero surgen problemas cuando la matriz $K^\top K$ está mal condicionada (tiene autovalores muy cercanos al 0), ya que entonces surgen oscilaciones grandes no

físicas en esos casos. En estos casos decimos que el sistema lineal $\mu = K\lambda$ está mal condicionado. Cuando esto sucede, histogramas verdaderos λ_1 y λ_2 que son muy diferentes pueden transformarse en histogramas alterados μ_1 y μ_2 que son muy similares. Como resultado, distinguir entre λ_1 y λ_2 basándose en datos alterados en el espacio S es muy difícil.

Para prevenir esto surgen las regularizaciones, mediante términos de penalización, o la detención prematura en métodos iterativos.

2.4. Regularización de Tikhonov

La Regularización de Tikhonov utiliza una función de regularización para controlar los resultados obtenidos. Se trata de estimar λ resolviendo:

$$\min_{\lambda \in \mathbb{R}^n} [(y - K\lambda)^\top (y - K\lambda) + \delta P(\lambda)]$$

El primer término como en el problema de mínimos cuadrados, el segundo penaliza las soluciones no físicas, y el coeficiente de regularización $\delta \geq 0$ regula la influencia de la regularización (vemos que si $\delta = 0$ se estaría resolviendo el problema de mínimos cuadrados del **Apartado 2.3** anterior). Algunas funciones de regularización son:

- **Curvatura:** $P(\lambda) = \|R\lambda\|^2$, donde R es un operador discretizado de la segunda derivada. Esto se puede calcular numéricamente, usando las derivadas por diferencias finitas:

$$f'(x) \approx \frac{f(x+h) - f(x-h)}{2h}$$

Y se llega a que la primera y segunda derivada discretas en un punto x_i son respectivamente: $x'_i = x_{i+1} - x_i$ y $x''_i = x_{i+1} - 2x_i + x_{i-1}$, y entonces la matriz R es:

$$R = \begin{pmatrix} -1 & 1 & 0 & \cdots & 0 & 0 \\ 1 & -2 & 1 & \cdots & 0 & 0 \\ 0 & 1 & -2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -2 & 1 \\ 0 & 0 & 0 & \cdots & 1 & -1 \end{pmatrix}$$

Ahora, nosotros en nuestro caso, añadiendo este término al MLE tenemos:

$$\min_{\lambda \in \mathbb{R}^n} [(y - K\lambda)^\top (y - K\lambda) + \delta \|R\lambda\|^2]$$

que por [2] se resuelve como $Uy = \lambda$ con:

$$U = (K^\top K + \delta R^\top R)^{-1} K^\top$$

- **SVD unfolding:** (Höcker and Kartvelishvili, 1996) [5]

$$P(\lambda) = \left\| R \begin{pmatrix} \frac{\lambda_1}{\lambda_1^{\text{MC}}} \\ \frac{\lambda_2}{\lambda_2^{\text{MC}}} \\ \vdots \\ \frac{\lambda_p}{\lambda_p^{\text{MC}}} \end{pmatrix} \right\|^2,$$

donde λ^{MC} es una referencia previa para λ con todas las componentes distintas de cero. Aquí el problema nos queda así:

$$\min_{\lambda \in \mathbb{R}^p} [(y - K\lambda)^\top (y - K\lambda) + \delta \|R(\lambda/\lambda^{\text{MC}})\|^2]$$

Y tomando

$$S = \begin{pmatrix} \frac{1}{\lambda_1^{\text{MC}}} & 0 & \cdots & 0 \\ 0 & \frac{1}{\lambda_2^{\text{MC}}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\lambda_n^{\text{MC}}} \end{pmatrix}$$

se puede linelizar así (en el **Apartado 4.2.** profundizaremos en los cálculos):

$$\hat{\lambda} = Uy = (K^\top K + \delta S^\top R^\top R S)^{-1} K^\top y$$

- **TUnfold:** (Schmitt, 2012) [6]:

$$P(\lambda) = \|R(\lambda - \lambda^{\text{MC}})\|^2$$

Aquí tenemos:

$$\min_{\lambda \in \mathbb{R}^p} [(y - K\lambda)^\top (y - K\lambda) + \delta \|R(\lambda - \lambda^{\text{MC}})\|^2]$$

Que por [2] se resuelve como:

$$\hat{\lambda} = (K^\top K + \delta(R^\top R))^{-1} (K^\top y + \delta(R^\top R)\lambda^{\text{MC}})$$

Lo que se puede escribir como

$$\hat{\lambda} = Uy + (K^\top K + \delta(R^\top R))^{-1} \delta(R^\top R)\lambda^{\text{MC}}$$

como transformación afín sobre y .

2.5. Método iterativo con 'early stopping'.

Si consideramos la función de verosimilitud, a pesar de los problemas que surgían cuando no se cumplían las hipótesis de la Proposición 2.2.2., el MLE siempre se puede encontrar computacionalmente utilizando el algoritmo esperanza-maximización (EM) (Dempster et al. (1977)) [7].

Este es un algoritmo iterativo es muy usado para encontrar soluciones de máxima verosimilitud en problemas con pocas observaciones.

Partiendo de algún valor inicial $\lambda^{(0)} > 0$, la iteración EM para el unfolding se da por: [2]

$$\lambda_j^{(k+1)} = \frac{\lambda_j^{(k)}}{\sum_{i=1}^n K_{ij}} \sum_{i=1}^n \frac{K_{ij} y_i}{\sum_{l=1}^n K_{il} \lambda_l^{(k)}}, \quad j = 1, \dots, p$$

La convergencia de esta iteración hacia un MLE (es decir, $\lambda^{(k)} \xrightarrow[k \rightarrow \infty]{} \hat{\lambda}_{\text{MLE}}$) fue probada por Vardi (Vardi et al. (1985)) [4]. Este algoritmo es conocido como la iteración D'Agostini.

Pero, desafortunadamente, el MLE en sí mismo es a menudo inútil. Debido a la naturaleza mal planteada del problema, exhibe grandes fluctuaciones no físicas. En otras palabras, la función de verosimilitud por sí sola no contiene suficiente información para restringir la solución. A medida que la iteración EM avanza, las soluciones suelen mejorarse primero, pero en algún punto comienzan a degradarse. Esto se debe a que el algoritmo empieza a sobreajustarse a las fluctuaciones de Poisson en y . Este comportamiento se puede aprovechar deteniendo la iteración antes de que comiencen a aparecer oscilaciones no físicas. El número de iteraciones k se convierte ahora en un parámetro de regularización que controla el equilibrio entre ajustar los datos y mitigar características no físicas.

2.6. Cálculo del sesgo, la varianza y la correlación

En el contexto del *unfolding*, es fundamental evaluar la precisión y la fiabilidad del estimador obtenido. Dos métricas clave para este análisis son el sesgo (*bias*) y la varianza (*variance*) del estimador.

El sesgo mide la diferencia sistemática entre el valor esperado del estimador y el valor verdadero de la cantidad que se quiere estimar. En otras palabras, el sesgo indica cuánto se desvía en promedio el estimador de la realidad. Formalmente, dado un estimador $\hat{\lambda}$ de un parámetro λ , el sesgo se define como:

$$\text{Bias}(\hat{\lambda}) = \mathbb{E}[\hat{\lambda}] - \lambda$$

Por otro lado, la varianza refleja la dispersión de las estimaciones en torno a su valor esperado. Una varianza alta indica que las estimaciones individuales fluctúan considerablemente, lo que puede dificultar la obtención de conclusiones precisas. La varianza del estimador se define como:

$$\text{Var}(\hat{\lambda}) = \mathbb{E} \left[(\hat{\lambda} - \mathbb{E}[\hat{\lambda}])^2 \right]$$

La covarianza en el contexto del *unfolding* mide cómo varían conjuntamente dos componentes del estimador de los espectros ajustados. Está dada por:

$$\text{Cov}(\hat{\lambda}_1, \hat{\lambda}_2) = E \left[(\hat{\lambda}_1 - E[\hat{\lambda}_1])(\hat{\lambda}_2 - E[\hat{\lambda}_2]) \right]$$

La correlación en el *unfolding* cuantifica la relación lineal estandarizada entre dos componentes ajustados del espectro y se define como:

$$\rho_{\hat{\lambda}_1 \hat{\lambda}_2} = \frac{\text{Cov}(\hat{\lambda}_1, \hat{\lambda}_2)}{\sqrt{\text{Var}(\hat{\lambda}_1) \cdot \text{Var}(\hat{\lambda}_2)}}$$

En el proceso de *unfolding*, la matriz de la aplicación lineal U y la matriz de respuesta K permiten expresar estos conceptos matricialmente. El sesgo del estimador se puede representar mediante la matriz B :

$$B = U \cdot K - I_n$$

Ya que:

$$\text{Bias}(\hat{\lambda}) = \mathbb{E}[\hat{\lambda}] - \lambda = Uy - \lambda = (UK - I_n)\lambda$$

Donde I_n es la matriz identidad de tamaño n . Esta expresión captura el desvío sistemático introducido por el método de *unfolding*.

Así mismo, la varianza del estimador se puede cuantificar mediante la siguiente expresión:

$$\text{Var} = U^\top U$$

Esta matriz proporciona una medida de la dispersión en las estimaciones obtenidas mediante el procedimiento de *unfolding*, cabe recordar que por construcción $\text{Var}(y) = 1$. Para la matriz de Correlación extraemos la diagonal de la matriz de varianza para formar la matriz diagonal D de varianzas:

$$D = \text{diag}(\text{Var}(e_1), \text{Var}(e_2), \dots, \text{Var}(e_n))$$

Donde $D_{ii} = \text{Var}(e_i) = \text{Var}_{ii}$ es la varianza del i -ésimo componente.

La matriz de correlación Corr se obtiene estandarizando la matriz de varianza. Primero calculamos la matriz $D^{-1/2}$, que es la raíz cuadrada inversa de D :

$$D^{-1/2} = \text{diag}\left(\frac{1}{\sqrt{\text{Var}(e_1)}}, \frac{1}{\sqrt{\text{Var}(e_2)}}, \dots, \frac{1}{\sqrt{\text{Var}(e_n)}}\right)$$

Luego, la matriz de correlación se calcula como:

$$\text{Corr} = D^{-1/2} \cdot \text{Var} \cdot D^{-1/2}$$

Por lo que dado el estimador $\hat{\lambda}$, el sesgo de $\hat{\lambda}$ viene dado por $\hat{\lambda} \cdot B \cdot \hat{\lambda}^\top$, la varianza viene dada por $\text{Var}(\hat{\lambda}) = \hat{\lambda} \cdot \text{Var} \cdot \hat{\lambda}^\top$. Y dados los estimadores $\hat{\lambda}_1$ y $\hat{\lambda}_2$, la covarianza de $\hat{\lambda}_1$ y $\hat{\lambda}_2$ es $\text{Cov}(\hat{\lambda}_1, \hat{\lambda}_2) = \hat{\lambda}_1 \cdot \text{Var} \cdot \hat{\lambda}_2^\top$, y la correlación entre $\hat{\lambda}_1$ y $\hat{\lambda}_2$ es $\rho_{\hat{\lambda}_1 \hat{\lambda}_2} = \hat{\lambda}_1 \cdot \text{Corr} \cdot \hat{\lambda}_2^\top$.

Estos métodos son utilizados para calcular el sesgo, las varianzas, las covarianzas y las correlaciones de los estimadores obtenidos mediante la regularización de Tikhonov con curvatura, el SVD Unfold y una iteración del método iterativo con 'early stopping', tras los cálculos detallados que se presentarán en el **Apartado 5.1** y para calcular la varianza de estimadores obtenidos por TUnfold, pero para el sesgo de del TUnfold y para más iteraciones del método D'Agostini no funciona. En el caso del TUnfold, ya que el estimador viene dado por

$$\hat{\lambda} = Uy + (K^\top K + \delta(R^\top R))^{-1} \delta(R^\top R) \lambda^{\text{MC}}$$

El término de la derecha es constante y no afecta a la varianza por lo que esta se calcula a partir de U como en los otros métodos, cabe remarcar que esta U es además idéntica a la de la regularización por la derivada. Sin embargo el término de la derecha sí que afecta al calculo del sesgo que se calculará explícitamente como:

$$\text{Bias}(\hat{\lambda}) = \hat{\lambda} - x_0 = Uy + (K^\top K + \delta(R^\top R))^{-1} \delta(R^\top R) \lambda^{\text{MC}} - x_0$$

Donde $y = Kx_0$.

Para el método D'Agostini en el que se hará uso de un x_0 inicial para calcular el sesgo explícitamente.

Para calcular el sesgo y la varianza a más iteraciones del algoritmo del algoritmo de D'Agostini lo representamos como:

$$\lambda_{\text{estimado}} = \mathbf{f}_n(\mathbf{y})$$

Donde excepto para $n = 1$, la f es difícil de obtener analíticamente, pero es posible calcularla iterativamente. Entonces, el sesgo de λ lo calculamos de la siguiente forma:

Partimos de un valor inicial real x_0 , y obtenemos un \mathbf{y} dado por $\mathbf{y} = \mathbf{K}x_0$. El sesgo será $x_0 - \mathbf{f}_n(\mathbf{K}x_0)$ calculado explícitamente.

Para calcular la varianza explícitamente es un poco más complicado, para obtener la varianza, y para que sea manejable, se puede hacer una aproximación lineal, que tiene sentido si la varianza de entrada (que tomaremos como 1, por la distribución $N(\mathbf{K}x, 1)$) como se verá en el **Apartado 3.1** es pequeña. Nos quedamos con el primer término del desarrollo en serie. Explicamos por qué funciona para una variable:

Supongamos que tenemos x que sigue $N(x_0, 1)$ y $y = f(x)$. Desarrollamos en serie de Taylor de primer orden y nos queda

$$y = f(x) \approx f(x_0) + f'(x_0)(x - x_0)$$

Son todas constantes salvo $f'(x_0)x$, por lo que sabemos que

$$y \sim N(f(x_0), f'(x_0))$$

donde hemos usado que $\sigma(x) = 1$, lo que justifica la aproximación.

Podemos calcular esto numéricamente, usando las derivadas por diferencias finitas que expusimos en el **Apartado 2.4**

$$f'(x) \approx \frac{f(x+h) - f(x-h)}{2h}$$

La idea es hacer esto coordenada a coordenada.

Para calcular la varianza de la coordenada i -ésima, se utiliza el siguiente algoritmo, modificando un poco arriba y abajo la componente i -ésima de $\boldsymbol{\mu}$. Algo como:

$$\sigma(f_n(Kx_0)) = \frac{f_n(Kx_0^+) - f_n(Kx_0^-)}{2h}$$

donde

$$\mathbf{x}_0^+ = \mathbf{x}_0 + (0, 0, 0, \dots, h, 0, 0, 0)$$

y

$$\mathbf{x}_0^- = \mathbf{x}_0 - (0, 0, 0, \dots, h, 0, 0, 0)$$

Esto sería σ , y la varianza es el cuadrado de esta cantidad.

2.7. Problemas y dificultades del *Unfolding*

Al realizar un *unfolding* sobre una medición, uno de los principales desafíos es la elección del parámetro de regularización, como δ en la regularización de Tikhonov o el número de iteraciones en el método de D'Agostini. Este parámetro controla la suavidad de la solución, influyendo directamente en la precisión del espectro estimado.

El principal problema radica en que este parámetro es difícil de determinar a priori, ya que su valor óptimo depende de los datos observados. Una elección incorrecta puede distorsionar el resultado: un valor demasiado alto suaviza excesivamente la solución, mientras que uno demasiado bajo permite que las fluctuaciones y errores dominen. Por ello, es esencial seleccionar este parámetro basándose en criterios objetivos y métodos empíricos que consideren la calidad del ajuste, como la minimización de discrepancias con simulaciones o el análisis de estabilidad del resultado obtenido, en el **Capítulo 4** se profundizará un poco más en este tema .

En el **Capítulo 2**, hemos detallado la formulación matemática del problema del unfolding y explorado técnicas clave como la regularización de Tikhonov y los mínimos cuadrados. Además, hemos introducido métodos para evaluar la precisión de las estimaciones mediante el cálculo del sesgo y la varianza.

Con esta base teórica, estamos preparados para aplicar estas técnicas a problemas específicos, analizando cómo se comportan en la práctica y optimizando los resultados. A partir de ahora, pondremos en práctica lo aprendido para evaluar su efectividad en casos concretos.

Capítulo 3

Presentación del problema

En el **Capítulo 3**, daremos el siguiente paso aplicando la teoría desarrollada en un caso práctico. Comenzaremos con la presentación de un problema específico de unfolding, donde abordaremos la elección y justificación de la matriz de dispersión K en nuestro caso, esencial para modelar correctamente la relación entre el espectro verdadero y el observado. A través de este análisis, exploraremos cómo los métodos de mínimos cuadrados pueden aplicarse directamente para resolver el problema, centrándonos en el impacto de la estructura de K en los resultados obtenidos. Este capítulo permitirá evaluar cómo las decisiones iniciales influyen en la precisión y estabilidad de las estimaciones.

3.1. Elección de K y Justificación

En nuestro caso consideramos K , que es una matriz de migración definida como:

$$K = \begin{pmatrix} 1-p & p & 0 & \cdots & 0 \\ p & 1-2p & p & \cdots & 0 \\ 0 & p & 1-2p & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & p & 1-p \end{pmatrix}$$

donde $y \in \mathbb{R}^n$, $\lambda \in \mathbb{R}^n$, $n \in \mathbb{N}$, y $p \in [0, 0,5]$.

En el problema general, K se suele calcular con simulación, lo que también puede generar distintos problemas de regularización o matrices mal condicionadas. La razón de nuestra elección de K es que buscamos algo manejable, general y que tenga sentido. Buscamos matrices de migración en las que solo se permite que haya migración al canal de al lado en función de una probabilidad p , con las que estudiaremos casos muy diagonales. Cabe remarcar que estamos quedándonos con los casos donde la dimensión de y es igual que la de λ (es decir, que la matriz K es cuadrada), lo que no siempre es así. Lo hacemos para simplificar el problema nuevamente, pero algunos cálculos matriciales se mantienen en los otros casos.

Además de esto, vamos a considerar $y \sim N(K\lambda, 1)$, ya que la normalización es adecuada cuando las cuentas son altas. Por lo tanto, es un problema particular pero sigue siendo bastante general.

Como hemos comentado antes, los problemas surgen cuando $K^\top K$ está mal condicionada. Por lo tanto, los problemas serán más pronunciados (y la regularización será más necesaria) en los valores de p donde la matriz $K^\top K$ tenga autovalores muy pequeños, lo cual puede llevar a inestabilidad numérica y a la necesidad de técnicas de regularización para asegurar soluciones estables y bien condicionadas. Todas las gráficas de este capítulo han sido obtenidas a partir de [12].

3.2. Ajuste por máxima verosimilitud

Si consideramos el caso 6×6 entonces tenemos la matriz

$$K = \begin{pmatrix} 1-p & p & 0 & 0 & 0 & 0 \\ p & 1-2p & p & 0 & 0 & 0 \\ 0 & p & 1-2p & p & 0 & 0 \\ 0 & 0 & p & 1-2p & p & 0 \\ 0 & 0 & 0 & p & 1-2p & p \\ 0 & 0 & 0 & 0 & p & 1-p \end{pmatrix}$$

Entonces, los valores de p en $[0, 0,5]$ para los que K tiene un autovalor 0 son:

$$p = 2 - \sqrt{3} \approx 0,268, \quad p = \frac{1}{3}, \quad \text{y} \quad p = \frac{1}{2}$$

Estos puntos valores son de vital importancia ya que es para estos valores para los que se hace necesaria la regularización como veremos más adelante y, al tomar $y \sim N(K\lambda, 1)$, el ajuste de máxima verosimilitud se convierte en el siguiente problema de mínimos cuadrados:

$$\min_{\lambda \in \mathbb{R}^6} \|K\lambda - y\|^2,$$

que se resuelve con la pseudoinversa como se expuso en el **Apartado 2.3**.

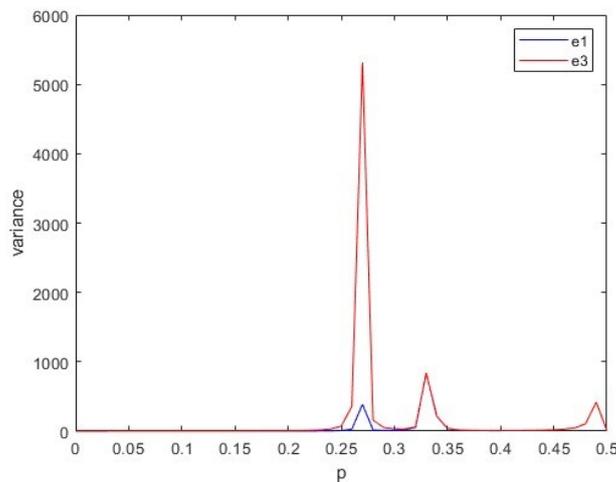


Figura 3.1: Varianza de los canales externo e_1 y central e_3 para $p \in [0, 0,5]$ en el caso 6×6 resuelto por mínimos cuadrados sin regularización.

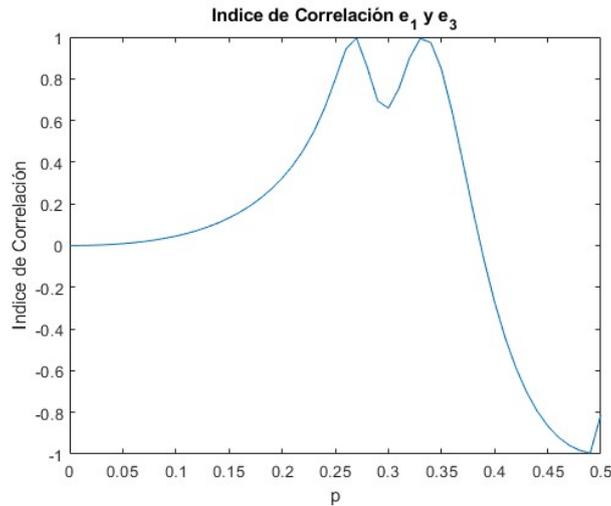


Figura 3.2: Índice de correlación entre el canal externo e_1 y el central e_3 para $p \in [0, 0,5]$ en el caso 6×6 resuelto por mínimos cuadrados sin regularización.

Como se puede observar en la Figura 3.1, la varianza de los estimadores se dispara para los valores de p que hacen que K tenga autovalores muy cercanos a 0. También es remarcable que el índice de correlación entre e_1 y e_3 es cercano a 1 cuando p se acerca a $p = 2 - \sqrt{3} \approx 0,268$ y $p = \frac{1}{3}$, y a -1 para valores de p cercanos a $p = \frac{1}{3}$, como se puede observar en la Figura 3.2. Esto significa que en los valores de p que hacen que K tenga autovalores muy cercanos a 0, la relación es casi lineal (positiva o negativa) perfecta entre ambos estimadores.

En este **Capítulo 3**, hemos aplicado la teoría de mínimos cuadrados a un caso práctico de unfolding. A través de la elección y análisis de la matriz de dispersión K , hemos demostrado cómo la estructura de K afecta significativamente los resultados obtenidos. Los cálculos realizados han permitido observar que, en problemas bien condicionados, los mínimos cuadrados proporcionan estimaciones precisas. Sin embargo, en casos donde la matriz K está mal condicionada, las soluciones pueden presentar inestabilidades.

Capítulo 4

Método de Tikhonov

En el **Capítulo 4**, nos enfocaremos en técnicas de regularización para mejorar la estabilidad de las soluciones en problemas mal condicionados de *unfolding*. En particular, exploraremos la regularización de Tikhonov y su aplicación en la resolución de sistemas lineales afectados por la presencia de autovalores pequeños en la matriz de dispersión K . Analizaremos cómo la regularización ayuda a controlar el equilibrio entre el sesgo y la varianza, y discutiremos la selección adecuada del coeficiente de regularización δ para obtener estimaciones más robustas y físicamente plausibles. Se ha hecho uso de [12] para la obtención de gráficas.

4.1. Curvatura

Tenemos

$$\min_{\lambda \in \mathbb{R}^n} [(y - K\lambda)^\top (y - K\lambda) + \delta \|R\lambda\|^2]$$

y podemos reescribir la función objetivo de la siguiente manera:

$$(y - K\lambda)^\top (y - K\lambda) + \delta \|R\lambda\|^2 = \|K\lambda - y\|^2 + \|\sqrt{\delta}R\lambda\|^2 = \left\| \begin{pmatrix} K \\ \sqrt{\delta}R \end{pmatrix} \lambda - \begin{pmatrix} y \\ 0 \end{pmatrix} \right\|^2$$

Que nos da un problema de mínimos cuadrados como el que expusimos antes, por lo que el minimizador viene dado por:

$$\hat{\lambda} = \begin{pmatrix} K \\ \sqrt{\delta}R \end{pmatrix}^\dagger \begin{pmatrix} y \\ 0 \end{pmatrix}$$

Como $\ker(K) \cap \ker(R) = \{0\}$, el minimizador es único y se puede simplificar de la siguiente manera:

$$\begin{aligned}
\hat{\lambda} &= \begin{pmatrix} K \\ \sqrt{\delta}R \end{pmatrix}^\dagger \begin{pmatrix} y \\ 0 \end{pmatrix} \\
&= \left(\begin{pmatrix} K \\ \sqrt{\delta}R \end{pmatrix}^\top \begin{pmatrix} K \\ \sqrt{\delta}R \end{pmatrix} \right)^{-1} \begin{pmatrix} K \\ \sqrt{\delta}R \end{pmatrix}^\top \begin{pmatrix} y \\ 0 \end{pmatrix} \\
&= \left((K^\top \quad \sqrt{\delta}R^\top) \begin{pmatrix} K \\ \sqrt{\delta}R \end{pmatrix} \right)^{-1} (K^\top \quad \sqrt{\delta}R^\top) \begin{pmatrix} y \\ 0 \end{pmatrix} \\
&= (K^\top K + \delta R^\top R)^{-1} K^\top y.
\end{aligned} \tag{4.1}$$

El problema es lineal sobre y dado un coeficiente de regularización δ tomando

$$U = (K^\top K + \delta R^\top R)^{-1} K^\top$$

Para $p = \frac{1}{3}$ y $\delta = 0, 0,1, 1$ salen las siguientes matrices U para el caso 3×3 .: (el valor $p = \frac{1}{3}$ hace que K sea mal condicionada en el caso 3×3)

$$U_1 = \begin{pmatrix} 0,50 & 0,50 & -1,50 \\ 1,50 & 0,00 & 2,50 \\ -1,00 & 0,50 & 0,00 \end{pmatrix}$$

$$U_2 = \begin{pmatrix} 0,95 & 0,33 & -0,28 \\ 0,33 & 0,33 & 0,33 \\ -0,28 & 0,33 & 0,95 \end{pmatrix}$$

$$U_3 = \begin{pmatrix} 0,56 & 0,33 & 0,10 \\ 0,33 & 0,33 & 0,33 \\ 0,10 & 0,33 & 0,56 \end{pmatrix}$$

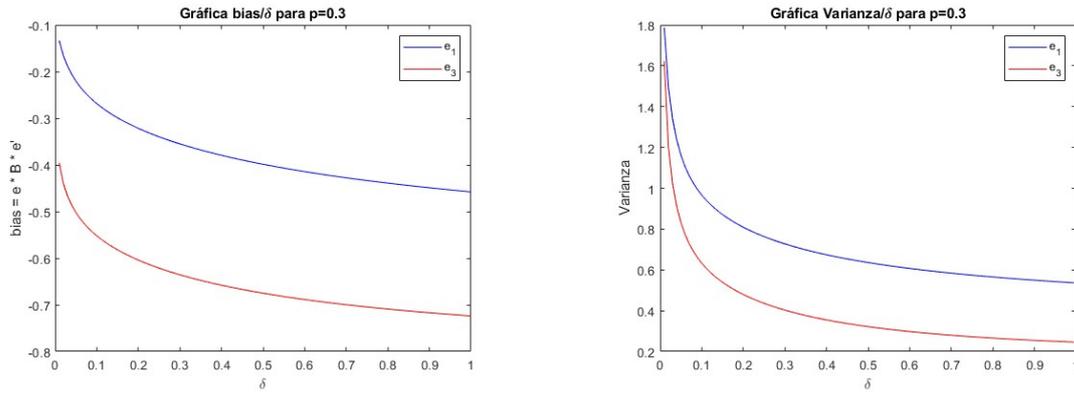
Como se puede observar, la matriz U se transforma para reflejar una mayor suavidad, eliminando entradas cercanas a cero o con valores negativos a medida que δ aumenta. Al incrementar δ , las matrices U pasan de tener valores más dispersos y, potencialmente, negativos, a tener valores más uniformes y positivos.

Volviendo al caso 6×6 anterior se puede observar como aumentando el valor de δ se reduce considerablemente la varianza, eso sí, se aumenta el sesgo como se puede ver en las siguientes Figuras 4.1a y 4.1b

4.2. SVD unfolding

Tomando

$$S = \begin{pmatrix} \frac{1}{\lambda_1^{MC}} & 0 & \cdots & 0 \\ 0 & \frac{1}{\lambda_2^{MC}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\lambda_n^{MC}} \end{pmatrix}$$



(a) Gráfica del sesgo

(b) Gráfica de la varianza

 Figura 4.1: Sesgo (*bias*) y la Varianza del canal externo e_1 y del central e_3 para $\delta \in [0,01, 1]$ en el caso 6×6 para el método de Tikhonov regularizado con la derivada.

como en el **Apartado 2.4**, se tiene que

$$P(\lambda) = \left\| R \begin{pmatrix} \frac{\lambda_1}{\lambda_{MC}^1} \\ \frac{\lambda_2}{\lambda_{MC}^2} \\ \vdots \\ \frac{\lambda_p}{\lambda_{MC}^p} \end{pmatrix} \right\|^2 = \|RS\lambda\|^2$$

Y entonces el problema de mínimos cuadrados queda tal que así:

$$\min_{\lambda \in \mathbb{R}^p} [(y - K\lambda)^\top (y - K\lambda) + \delta \|RS\lambda\|^2]$$

que con un razonamiento parecido al de antes se llega a

$$\hat{\lambda} = \begin{pmatrix} K \\ \sqrt{\delta}RS \end{pmatrix}^\dagger \begin{pmatrix} y \\ 0 \end{pmatrix} = (K^\top K + \delta S^\top R^\top RS)^{-1} K^\top y$$

Lo que nos lleva a:

$$U = (K^\top K + \delta S^\top R^\top RS)^{-1} K^\top$$

Para $p = \frac{1}{3}$ y $\delta = 0, 0,1, 1$ salen las siguientes matrices U para el caso 3×3 usando $\lambda_{MC} = [1, 2, 1]$:

$$U_1 = \begin{pmatrix} 0,50 & 0,50 & -1,50 \\ 1,50 & 0,00 & 2,50 \\ -1,00 & 0,50 & 0,00 \end{pmatrix}$$

$$U_2 = \begin{pmatrix} 0,86 & 0,25 & -0,36 \\ 0,50 & 0,50 & 0,50 \\ -0,36 & 0,25 & 0,86 \end{pmatrix}$$

$$U_3 = \begin{pmatrix} 0,48 & 0,25 & 0,02 \\ 0,50 & 0,50 & 0,50 \\ 0,02 & 0,25 & 0,48 \end{pmatrix}$$

Y si usamos $\lambda_{MC} = 10 * [1, 2, 1]$, salen:

$$U_1 = \begin{pmatrix} 0,50 & 0,50 & -1,50 \\ 1,50 & 0,00 & 2,50 \\ -1,00 & 0,50 & 0,00 \end{pmatrix}$$

$$U_2 = \begin{pmatrix} 1,00 & 0,25 & -0,50 \\ 0,50 & 0,50 & 0,50 \\ -0,50 & 0,25 & 1,00 \end{pmatrix}$$

$$U_3 = \begin{pmatrix} 0,98 & 0,25 & -0,48 \\ 0,50 & 0,50 & 0,50 \\ -0,48 & 0,25 & 0,98 \end{pmatrix}$$

Aquí, de la misma forma que sucedía cuando se regularizaba con la derivada, la matriz U se va suavizando a medida que δ aumenta. Es notable que se suaviza más “rápido” (a menor δ) con la primera referencia $\lambda_{MC} = [1, 2, 1]$ que con la segunda $\lambda_{MC} = 10 * [1, 2, 1]$ que tiene más oscilaciones y es una referencia menos acertada. Por lo que en ese caso habrá que considerar tomar un δ mayor para solventar la mala elección de λ_{MC} teniendo en cuenta como afectará esto a la varianza.

En las siguientes Figuras 4.2 observamos lo explicado aquí, se nota como para todo $p \in [0, 0,5]$ para los mismos coeficientes de regularización $\delta = 0,1$ y $\delta = 1$ con $\lambda_{MC} = [1, 1, 4, 5, 2, 1]$ el sesgo es mayor, pero la varianza (sobretudo para e_3) y la correlación entre e_1 y e_3 son menores que para $\lambda_{MC} = 10 * [1, 1, 4, 5, 2, 1]$

4.3. TUnfold

Aquí tenemos:

$$\min_{\lambda \in \mathbb{R}^p} [(y - K\lambda)^\top (y - K\lambda) + \delta \|R(\lambda - \lambda^{MC})\|^2]$$

Haciendo los cálculos como antes se llega a:

$$\begin{aligned} (y - K\lambda)^\top (y - K\lambda) + \delta \|R(\lambda - \lambda^{MC})\|^2 \\ = \|K\lambda - y\|^2 + \delta \|R(\lambda - \lambda^{MC})\|^2 \\ = \left\| \begin{pmatrix} K \\ \sqrt{\delta}R \end{pmatrix} \lambda - \begin{pmatrix} y \\ \sqrt{\delta}R\lambda^{MC} \end{pmatrix} \right\|^2 \end{aligned} \quad (4.2)$$

Que nos da:

$$\hat{\lambda} = \begin{pmatrix} K \\ \sqrt{\delta}R \end{pmatrix}^\dagger \begin{pmatrix} y \\ \sqrt{\delta}R\lambda^{MC} \end{pmatrix} = (K^\top K + \delta(R^\top R))^{-1} (K^\top y + \delta(R^\top R)\lambda^{MC})$$

La transformación afín discutida en el **Apartado 2.4** indica que la varianza de los estimadores obtenidos mediante el método TUnfold es idéntica a la obtenida al regularizar por la curvatura (como se comentó en **Apartado 2.6**). No obstante, TUnfold introduce un sesgo hacia el vector λ^{MC} , lo que permite reducir el sesgo

en comparación con el método de la derivada. Este sesgo dirigido hacia λ^{MC} puede mejorar la precisión de la estimación, siempre que λ^{MC} se elija adecuadamente. Por tanto, una buena elección de λ^{MC} es crucial, ya que permite minimizar tanto el sesgo como la varianza, optimizando la calidad del estimador final. En el siguiente **Apartado 4.4** se profundizará más en esto.

4.4. Elección del coeficiente de regularización δ .

Como hemos comentado previamente, el coeficiente de regularización δ es fundamental en el unfolding de Tikhonov. En problemas de unfolding, que son típicamente mal planteados, la regularización mediante δ equilibra el ajuste a los datos con la estabilidad de la solución. Específicamente, δ controla la suavidad del estimador, previniendo oscilaciones no físicas y filtrando el ruido en los datos.

La selección adecuada de δ se basa en minimizar la función de intercambio sesgo-varianza (*Bias/Variance Trade Off*), que mide la suma del sesgo al cuadrado y la varianza de todos los canales del estimador, en función de δ . Un δ más alto introduce mayor suavidad (aumentando el sesgo) pero reduce la varianza al suprimir fluctuaciones. El valor óptimo de δ es aquel que minimiza esta función, logrando un equilibrio entre precisión y estabilidad, lo que resulta en una solución robusta y fiable. Este enfoque garantiza que el *Unfolding* de los datos observados sea lo suficientemente suave para ser físicamente plausible, pero lo suficientemente ajustado a los datos para minimizar la incertidumbre en las predicciones.

En el caso 6×6 que consideramos antes, si hacemos *Unfolding* por Tikhonov regularizado por la derivada, los valores de δ que minimizan la función de intercambio sesgo-varianza son $\delta \approx 0,58$ y $\delta \approx 3,18$ para $p = 0,1$ y $p = 0,3$ respectivamente como observamos en la Figura 4.3. Esto nos lleva a pensar que lo que se venía intuyendo, según p se acerque a los valores que hacen K mal condicionada será más necesaria la regularización, por ende, mayor coeficiente de regularización δ .

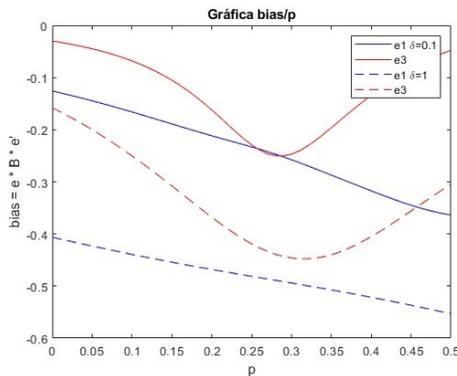
En las Figuras 4.4 y 4.5 se puede observar que si lo hacemos por SVD con $\lambda_{MC} = [1, 1, 4, 5, 2, 1]$ los valores de δ son $\delta \approx 2,76$ y $\delta \approx 11,68$ para $p = 0,1$ y $p = 0,3$ respectivamente; y si lo hacemos $\lambda_{MC} = 10 * [1, 1, 4, 5, 2, 1]$ sale $\delta \geq 100$ para $p = 0,1$ y $p = 0,3$. Esto resalta, la importancia de hacer una buena elección de λ_{MC} ya que una mala elección dispara la varianza y hace necesario un coeficiente de regularización muy grande.

En las Figuras 4.6, 4.7 y 4.8, observamos como para distintas matrices M (con el arreglo $\lambda_{MC} = Mx_0$) nos salen distintos valores δ que minimizan la función de sesgo-varianza para el caso TUnfold. En la Figura 4.6 el valor de δ tiende a infinito ya que el λ_{MC} que estamos considerando es exactamente x_0 por lo que, cuanto menos se consideren los datos observados y más precisión se consigue. En el segundo caso estaríamos tomando $\lambda_{MC} = [1, 1, 1, 1, 1, 1]$, con este vector los valores de δ que minimizan son $\delta \approx 27,64$ y $\delta \approx 31,52$ para $p = 0,1$ y $p = 0,3$ respectivamente. Para el último caso, tomando un vector desescalado $\lambda_{MC} = 10 * [1, 1, 1, 1, 1, 1]$ los valores de δ que minimizan quedan $\delta \approx 0,019$ y $\delta \approx 0,022$ en los respectivos valores de p , aquí lo que sucede es que el valor de δ disminuye para regular la desproporción de λ_{MC} . Sería interesante estandarizar λ_{MC} , ya que nos daría una visión más clara de como afecta el coeficiente de regularización δ lo que nos permite compararlo más

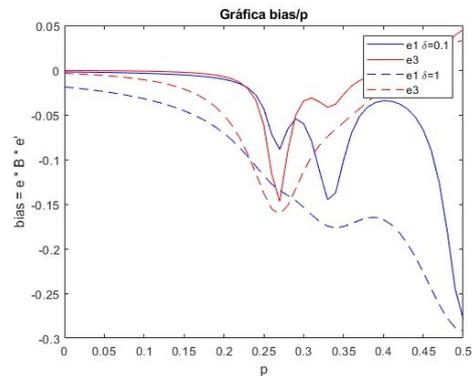
fácilmente con otros métodos.

En este **Capítulo 4**, hemos explorado la regularización de Tikhonov como una herramienta clave para mejorar la estabilidad en problemas de *unfolding* mal condicionados. A través de su aplicación, hemos visto cómo el parámetro de regularización δ permite controlar el equilibrio entre sesgo y varianza, logrando soluciones más estables y evitando oscilaciones no físicas en los resultados. Los análisis realizados subrayan la importancia de seleccionar adecuadamente δ para obtener estimaciones fiables, destacando la utilidad de la regularización en situaciones donde los métodos sin regularización fallan.

Sesgo (λ_{MC})

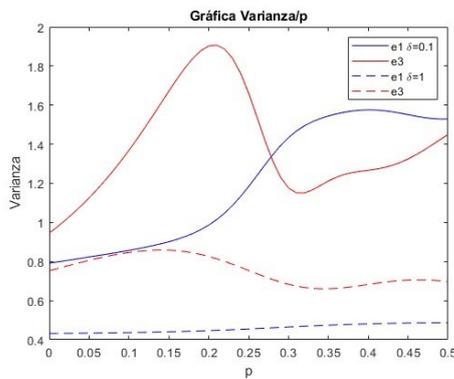


(a) $\lambda_{MC} = [1, 1, 4, 5, 2, 1]$

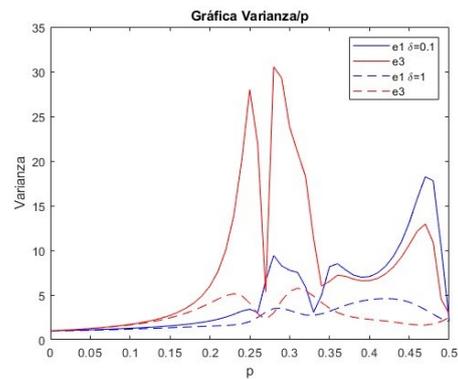


(b) $\lambda_{MC} = 10 * [1, 1, 4, 5, 2, 1]$

Varianza (λ_{MC})

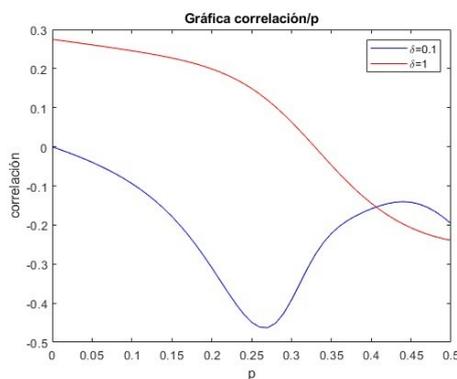


(c) $\lambda_{MC} = [1, 1, 4, 5, 2, 1]$

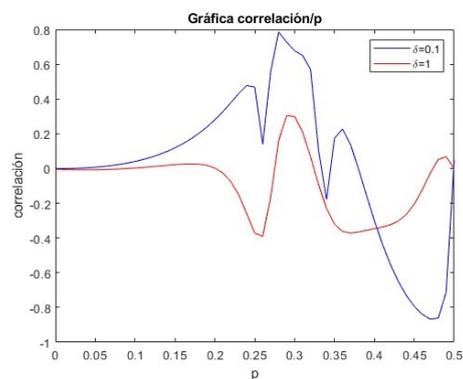


(d) $\lambda_{MC} = 10 * [1, 1, 4, 5, 2, 1]$

Correlación (λ_{MC})



(e) $\lambda_{MC} = [1, 1, 4, 5, 2, 1]$



(f) $\lambda_{MC} = 10 * [1, 1, 4, 5, 2, 1]$

Figura 4.2: Sesgo (*bias*) y varianza del canal externo e_1 y del central e_3 , y del índice de correlación de e_1 con e_3 para $p \in [0, 0,5]$ en el caso 6×6 para el método de Tikhonov regularizado con SVD para $\lambda_{MC} = [1, 1, 4, 5, 2, 1]$ y para $\lambda_{MC} = 10 * [1, 1, 4, 5, 2, 1]$.

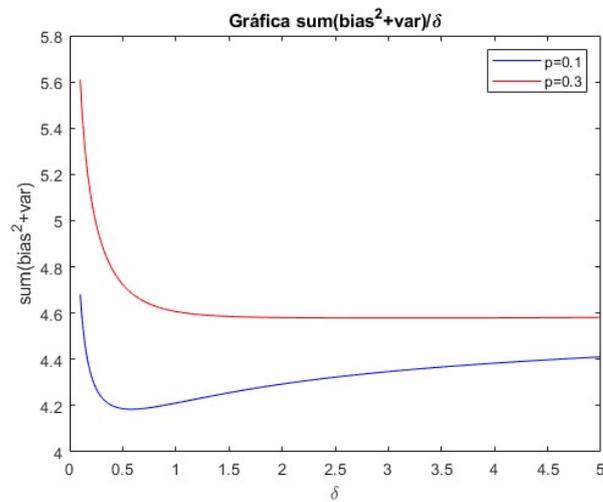


Figura 4.3: Suma del cuadrado del sesgo más la varianza para $\delta \in [0,1,5]$, $p = 0,1$ y $p = 0,3$ en el caso 6×6 con el método de Tikhonov regularizado con la derivada.

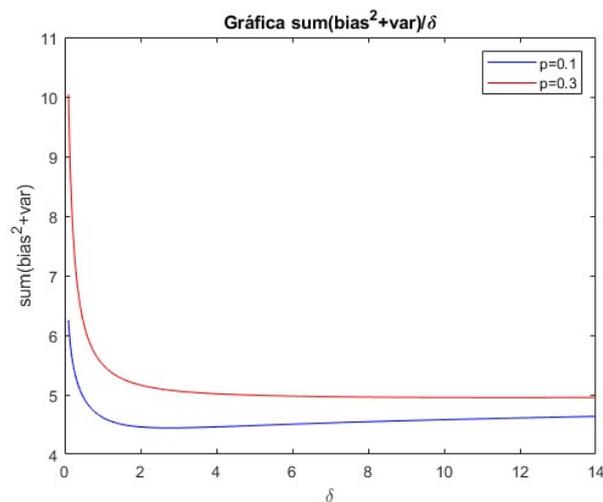


Figura 4.4: Suma del cuadrado del sesgo más la varianza para $\delta \in [0,1,14]$, $p = 0,1$ y $p = 0,3$ en el caso 6×6 con el método SVD para $\lambda_{MC} = [1, 1, 4, 5, 2, 1]$.

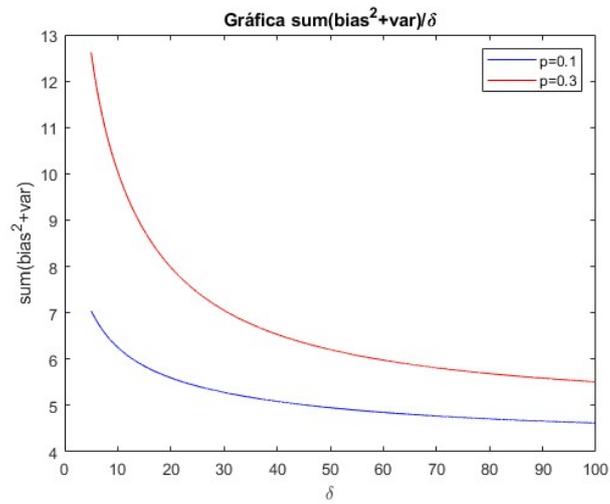


Figura 4.5: Suma del cuadrado del sesgo más la varianza para $\delta \in [5, 100]$, $p = 0,1$ y $p = 0,3$ en el caso 6×6 con el método SVD para $\lambda_{MC} = 10 * [1, 1, 4, 5, 2, 1]$.

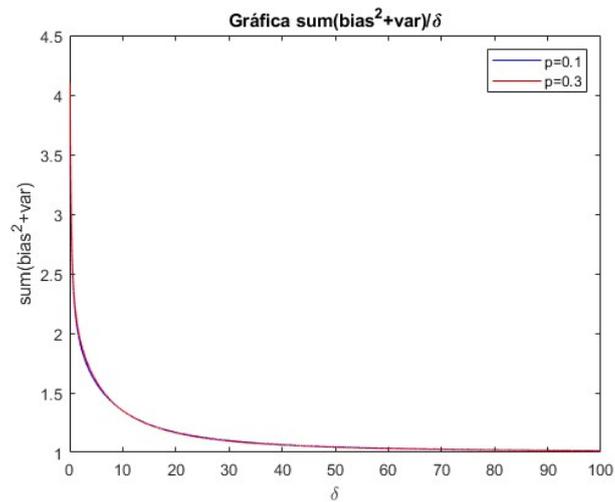


Figura 4.6: Suma del cuadrado del sesgo más la varianza para $\delta \in [0,1, 100]$, $p = 0,1$ y $p = 0,3$ en el caso 6×6 con el método TUnfold para $x_0 = [1, 2, 3, 4, 5, 6]$ y $\lambda_{MC} = Id * x_0$.

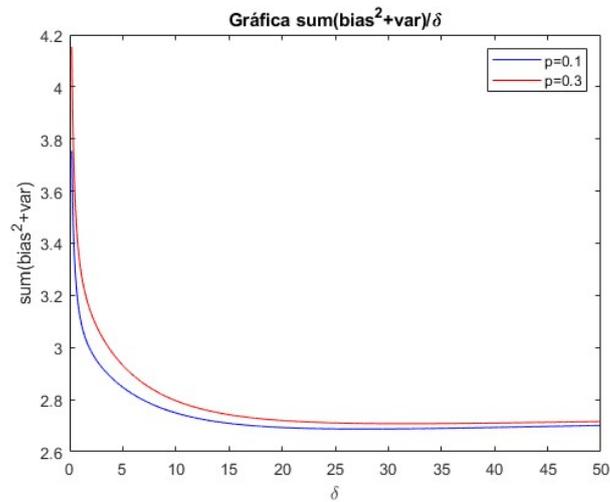
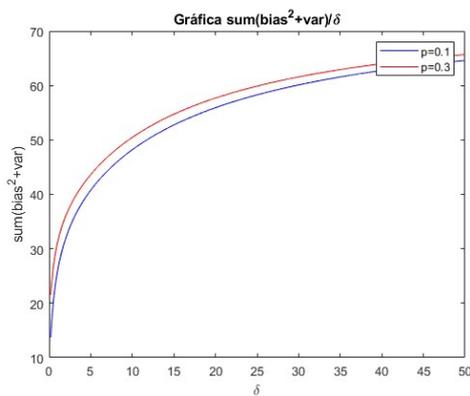
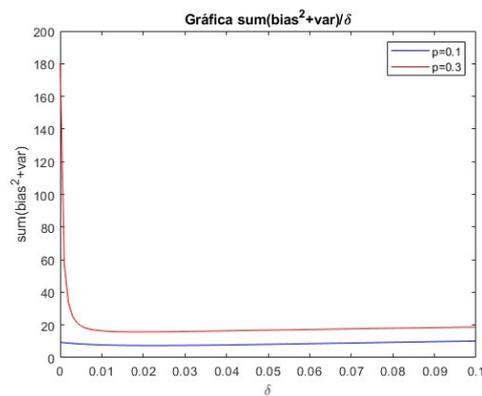


Figura 4.7: Esta gráfica muestra el la suma del cuadrado del sesgo más la varianza para $\delta \in [0, 50]$, $p = 0,1$ y $p = 0,3$ en el caso 6×6 con el método TUnfold para $x_0 = [1, 2, 3, 4, 5,]$ y $\lambda_{MC} = \text{diag}([1, 1/2, 1/3, 1/4, 1/5, 1/6]) * x_0$.



(a) $\delta \in [0, 50]$



(b) $\delta \in [0, 0,1]$

Figura 4.8: Suma del cuadrado del sesgo más la varianza para $\delta \in [0, 0,1]$, $p = 0,1$ y $p = 0,3$ en el caso 6×6 con el método TUnfold para $x_0 = [1, 2, 3, 4, 5, 6]$ y $\lambda_{MC} = 10 * \text{diag}([1, 1/2, 1/3, 1/4, 1/5, 1/6]) * x_0$

Capítulo 5

Método iterativo con 'early stopping'

En el *unfolding* estadístico, donde se busca recuperar distribuciones verdaderas a partir de datos observados distorsionados, el método iterativo con 'early stopping' se destaca como una técnica fundamental para evitar el sobreajuste. Este capítulo se divide en dos partes principales. En la primera, se examina el comportamiento del método tras una única iteración, comparándolo con otros enfoques como la regularización de Tikhonov, para evaluar su eficacia inicial. En la segunda parte, se analiza el efecto de realizar múltiples iteraciones, destacando cómo la elección del número óptimo de iteraciones es crucial para balancear adecuadamente el sesgo y la varianza. Este análisis detallado subraya la importancia de detener las iteraciones en el momento adecuado para obtener estimaciones precisas y estables en problemas mal condicionados. Se ha usado [12] para obtener las gráficas de este capítulos.

5.1. Una iteración

Cada iteración del algoritmo D'Agostini se puede expresar como lineal (como hicimos con la regularización de Tikhonov), dado un $\lambda^{(k)}$, de esta manera:

$$\begin{aligned} \begin{pmatrix} \lambda_1^{(k+1)} \\ \vdots \\ \lambda_n^{(k+1)} \end{pmatrix} &= \begin{pmatrix} \frac{\lambda_1^{(k)}}{\sum_{i=1}^n K_{i1}} \sum_{i=1}^n \frac{K_{i1} y_i}{\sum_{l=1}^n K_{il} \lambda_l^{(k)}} \\ \vdots \\ \frac{\lambda_n^{(k)}}{\sum_{i=1}^n K_{in}} \sum_{i=1}^n \frac{K_{in} y_i}{\sum_{l=1}^n K_{il} \lambda_l^{(k)}} \end{pmatrix} = \\ &= \begin{pmatrix} \frac{\lambda_1^{(k)}}{\sum_{i=1}^n K_{i1}} & 0 & \cdots & 0 \\ 0 & \frac{\lambda_2^{(k)}}{\sum_{i=1}^n K_{i2}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{\lambda_n^{(k)}}{\sum_{i=1}^n K_{in}} \end{pmatrix} \begin{pmatrix} \frac{K_{11}}{\sum_{l=1}^n K_{1l} \lambda_l^{(k)}} & \cdots & \frac{K_{n1}}{\sum_{l=1}^n K_{nl} \lambda_l^{(k)}} \\ \vdots & \ddots & \vdots \\ \frac{K_{1n}}{\sum_{l=1}^n K_{1l} \lambda_l^{(k)}} & \cdots & \frac{K_{nn}}{\sum_{l=1}^n K_{nl} \lambda_l^{(k)}} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \end{aligned}$$

Que es igual a:

$$\begin{pmatrix} \frac{\lambda_1^{(k)}}{\sum_{i=1}^n K_{i1}} & 0 & \cdots & 0 \\ 0 & \frac{\lambda_2^{(k)}}{\sum_{i=1}^n K_{i2}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{\lambda_n^{(k)}}{\sum_{i=1}^n K_{in}} \end{pmatrix} K^\top \begin{pmatrix} \frac{1}{\sum_{l=1}^n K_{1l}\lambda_l^{(k)}} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sum_{l=1}^n K_{2l}\lambda_l^{(k)}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sum_{l=1}^n K_{nl}\lambda_l^{(k)}} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

Lo que haciendo

$$P = \begin{pmatrix} \frac{\lambda_1^{(k)}}{\sum_{i=1}^n K_{i1}} & 0 & 0 & \cdots & 0 \\ 0 & \frac{\lambda_2^{(k)}}{\sum_{i=1}^n K_{i2}} & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & \frac{\lambda_n^{(k)}}{\sum_{i=1}^n K_{in}} \end{pmatrix}$$

y

$$Q = \begin{pmatrix} \frac{1}{\sum_{l=1}^n K_{1l}\lambda_l^{(k)}} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sum_{l=1}^n K_{2l}\lambda_l^{(k)}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sum_{l=1}^n K_{nl}\lambda_l^{(k)}} \end{pmatrix}$$

Nos queda tal que así:

$$\lambda^{(k+1)} = PK^\top Qy$$

Por lo tanto, $U = PK^\top Q$ es la matriz asociada a la aplicación lineal de una iteración del algoritmo de D'Agostini. Cabe remarcar que las matrices P y Q son dependientes del λ_k que se use a una iteración, por lo que no es posible linealizarlo así para más iteraciones. Podemos calcular la matriz del sesgo B y la de varianza Var a partir de U , lo que nos permite comparar una iteración de este método con los de Tikhonov. Por ejemplo, si tomamos $\lambda_0 = [1, 1, 1, 1, 1, 1]$ aumenta el sesgo y disminuye la varianza de e_3 y disminuye el sesgo y aumenta la varianza de e_1 en relación con lo que sucede si usamos $\lambda_0 = [1, 1, 4, 5, 2, 1]$ que tiene más oscilaciones como se puede observar en la Figuras 5.1 y 5.2

5.2. Más iteraciones

Para más iteraciones, no se puede linealizar por lo que la varianza y el sesgo los obtenemos para un x_0 real con el que calculamos el respectivo y alterado y calculamos la varianza y el sesgo como se comentó en los **Apartados 2.5 y 2.6**. Si consideramos el caso 6×6 y tomamos $x_0 = [5, 4, 7, 2, 8, 1]$, λ_i es el respectivo a 10^{i-1} iteraciones del algoritmo para $i = 1, 2, 3, 4$ partiendo de una estimación inicial plana $\lambda_0 = [1, 1, 1, 1, 1, 1]$ y otra con más oscilaciones $\lambda_0 = [1, 4, 8, 8, 4, 1]$, para casos con menos $p = 0,1$ y más incertidumbre $p = 0,26$ (el segundo cercano al punto crítico: $p \approx 0,268$).

$$\lambda_0 = [1, 1, 1, 1, 1, 1]$$

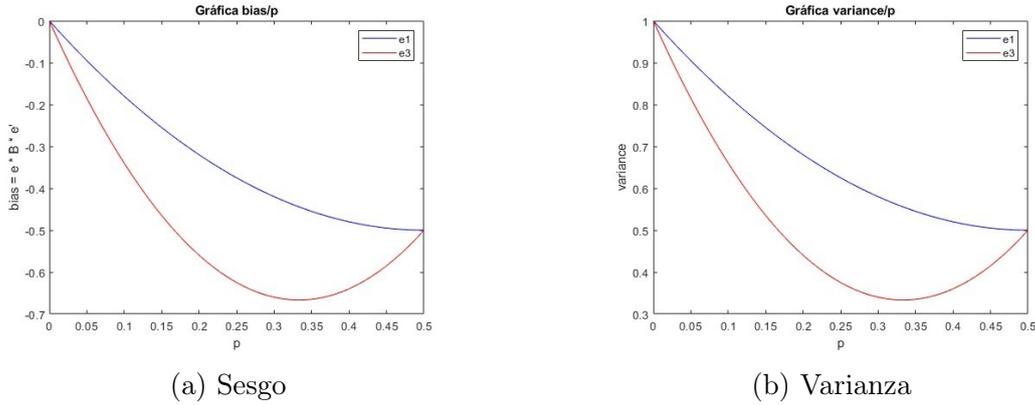


Figura 5.1: Sesgo y varianza de e_1 y e_3 para una iteración del algoritmo partiendo de $\lambda_0 = [1, 1, 1, 1, 1, 1]$ con $p \in [0, 0,5]$.

$$\lambda_0 = [1, 1, 4, 5, 2, 1]$$

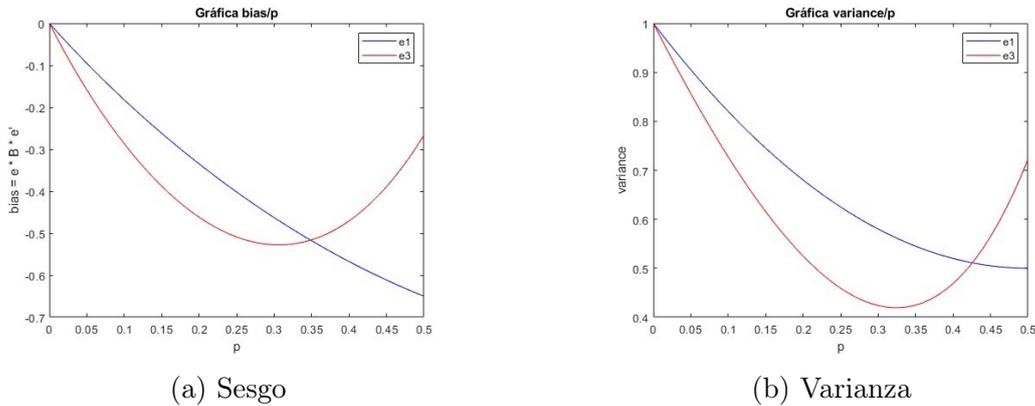


Figura 5.2: Sesgo y varianza de e_1 y e_3 para una iteración del algoritmo partiendo de $\lambda_0 = [1, 1, 4, 5, 2, 1]$ con $p \in [0, 0,5]$.

- Poca incertidumbre $p = 0,1$: Como se puede observar en las Figuras 5.3 y 5.4, en los casos con poca incertidumbre, donde la matriz K está bien condicionada, independientemente de usar un λ_0 más plano o con más oscilaciones el estimador λ es casi igual a x_0 a partir de las 10 iteraciones del proceso, el sesgo en ambos casos se hace casi nulo y la varianza queda igual a 1, que es la varianza de x_0 .

Como habíamos comentado previamente en el **Apartado 2.2**, cuando la matriz K está bien condicionada y es cuadrada con $K^{-1}y \geq 0$ el algoritmo obtiene el estimador de máxima verosimilitud logarítmica λ_{MLE} , que proporciona una buena estimación.

- Mucha incertidumbre $p = 0,26$: En este caso no sucede como con $p = 0,1$ en el sentido de que el sesgo de los estimadores no se anula en menos de mil iteraciones como se puede observar en las Figuras 5.5 y 5.6. Es notable también que el sesgo se mantiene mayor y la varianza se mantiene menor en los canales

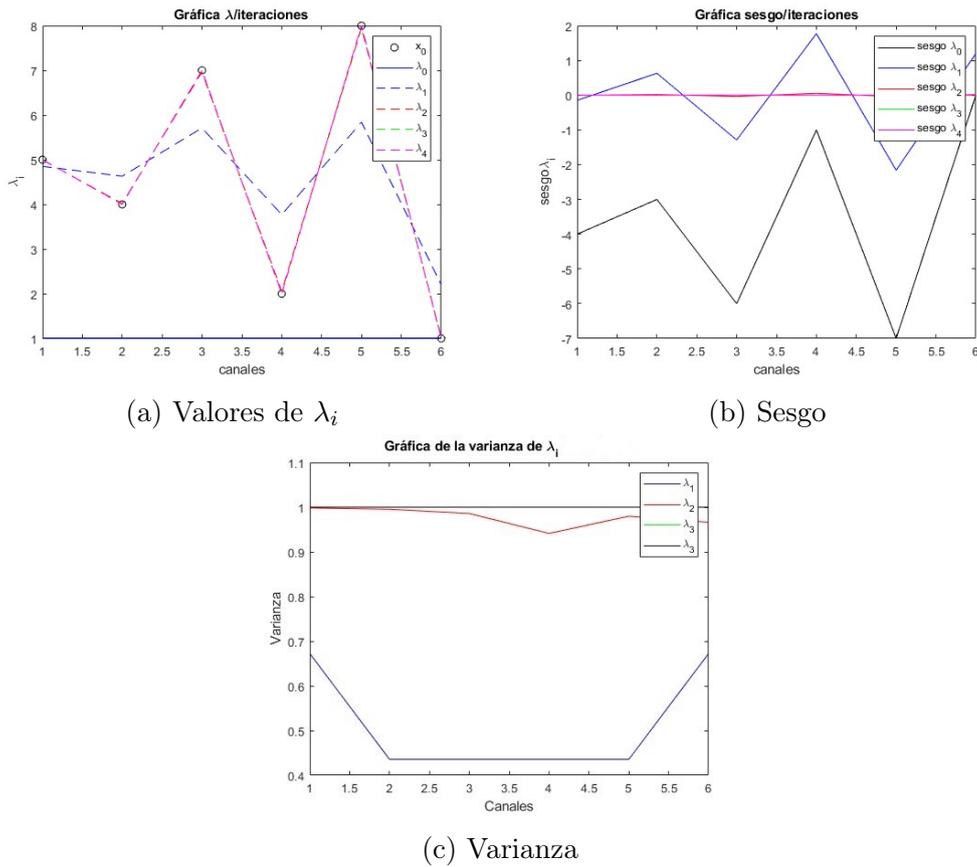
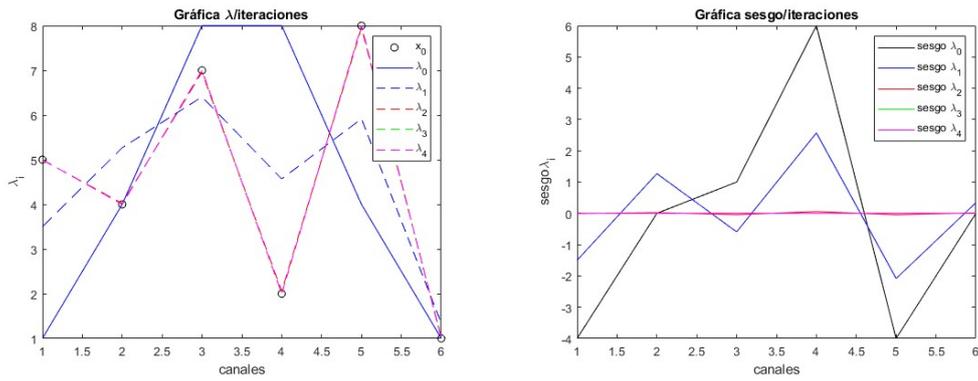


Figura 5.3: Gráficas con los valores, el sesgo y la varianza de λ_i para 1, 10, 100, 1000 iteraciones del método iterativo con 'early stopping' con $p = 0,1$ partiendo de $\lambda_0 = [1, 1, 1, 1, 1]$

para los que la estimación inicial λ_0 es menos precisa que en los canales donde sí que es más acertada.

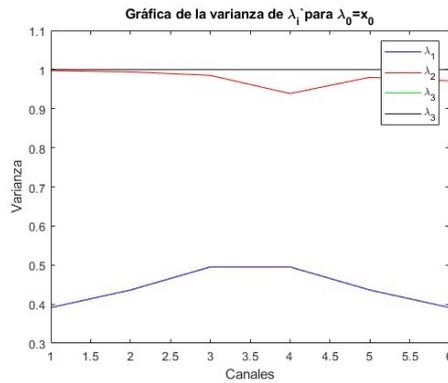
Elegir el número de iteraciones en el algoritmo de D'Agostini es clave para equilibrar el ajuste adecuado de los datos y evitar el sobreajuste. En las primeras iteraciones, el sesgo es alto y la varianza baja, pero a medida que aumentan, el sesgo disminuye y la varianza aumenta, acercando la solución a los datos. Sin embargo, demasiadas iteraciones pueden llevar a un sobreajuste, donde el modelo ajusta el ruido en lugar de la distribución real. La parada temprana ('early stopping') es esencial para detener el proceso antes de este punto, logrando un balance óptimo entre sesgo y varianza (*Bias/Variance Trade Off*). La selección precisa del número de iteraciones se ajusta mediante simulación por distintos métodos.

En este capítulo, hemos explorado el método iterativo con 'early stopping' como una estrategia clave en el unfolding estadístico. A través del análisis de una única iteración, hemos demostrado cómo este método se compara favorablemente con otros enfoques, como la regularización de Tikhonov, en términos de control del sesgo y la varianza. Al extender el análisis a múltiples iteraciones, se ha demostrado que, cuando el problema está bien condicionado, el estimador de máxima verosimilitud es una buena estimación por sí mismo, lo que hace innecesario el uso de 'early stopping'.



(a) Valores de λ_i

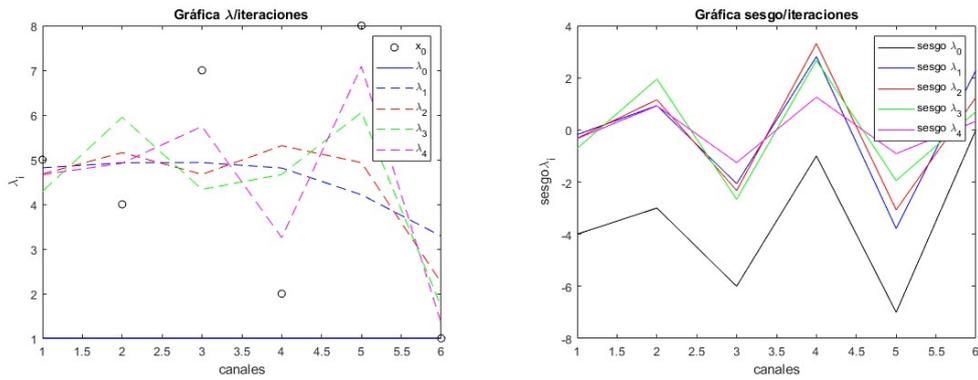
(b) Sesgo



(c) Varianza

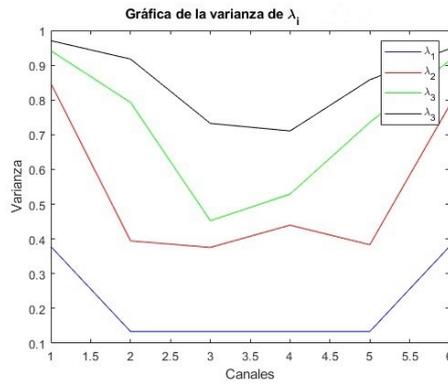
Figura 5.4: Gráficas con los valores, el sesgo y la varianza de λ_i para 1, 10, 100, 1000 iteraciones del método iterativo con 'early stopping' con $p = 0,1$ partiendo de $\lambda_0 = [1, 4, 8, 8, 4, 1]$

Sin embargo, en situaciones donde el problema es mal condicionado, detener las iteraciones en el momento adecuado se vuelve esencial para evitar el sobreajuste y garantizar estimaciones robustas.



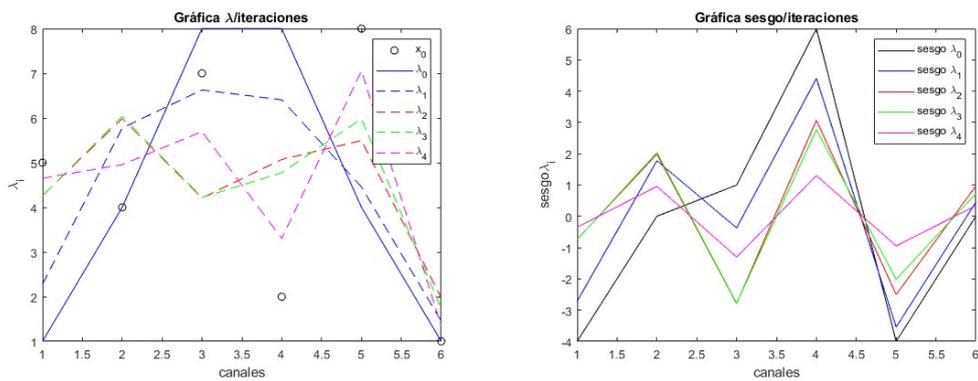
(a) Valores de λ_i

(b) Sesgo



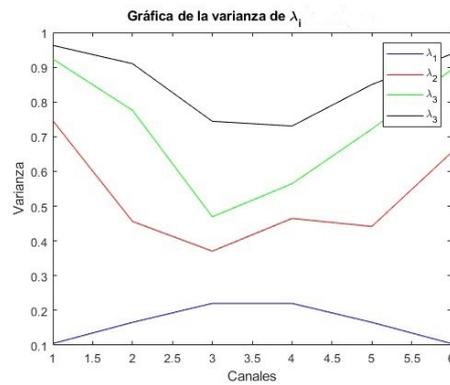
(c) Varianza

Figura 5.5: Gráficas con los valores, el sesgo y la varianza de λ_i para 1, 10, 100, 1000 iteraciones del método iterativo con 'early stopping' con $p = 0,26$ partiendo de $\lambda_0 = [1, 1, 1, 1, 1, 1]$



(a) Valores de λ_i

(b) Sesgo



(c) Varianza

Figura 5.6: Gráficas con los valores, el sesgo y la varianza de λ_i para 1, 10, 100, 1000 iteraciones del método iterativo con 'early stopping' con $p = 0,26$ partiendo de $\lambda_0 = [1, 4, 8, 8, 4, 1]$

Capítulo 6

Conclusiones

En este trabajo, hemos abordado el problema de *unfolding* en estadística, centrándonos en la estimación de distribuciones verdaderas a partir de observaciones distorsionadas. A lo largo del estudio, se han explorado diversas metodologías, como los mínimos cuadrados, la regularización de Tikhonov, haciendo énfasis en la regularización con la curvatura, el enfoque de SVD y TUnfold, y el método iterativo con 'early stopping'. Cada técnica ha demostrado sus ventajas y limitaciones en función de la condición del problema y la calidad de los datos.

En particular, se ha resaltado la importancia de la elección adecuada del coeficiente de regularización y del número de iteraciones en los métodos iterativos. Estos parámetros son cruciales para equilibrar el sesgo y la varianza, evitando tanto el sobreajuste como la pérdida de información relevante. Recomendamos utilizar métodos que introducen regularización, como Tikhonov o *early stopping*, cuando el problema está mal condicionado, es decir, cuando la matriz de dispersión es mal condicionada o tiene autovalores muy cercanos al cero. En estos casos, con la regularización se permite introducir sesgo a los resultados para limitar la varianza y obtener una solución más estable y física.

Por otro lado, cuando el problema está bien condicionado y se puede resolver por mínimos cuadrados directamente, no recomendamos someter la solución a un sesgo innecesario. En estas situaciones, al no haber una varianza significativa, los mínimos cuadrados ofrecen una solución precisa sin necesidad de regularización adicional.

En resumen, la aplicación de estos métodos en el problema del *unfolding* ha permitido una comprensión más profunda de los desafíos y las técnicas disponibles para abordar problemas inversos mal planteados. Los resultados obtenidos subrayan la necesidad de un enfoque cuidadoso y adaptado a las características específicas de cada caso, garantizando estimaciones precisas y estables en el contexto de mediciones experimentales.

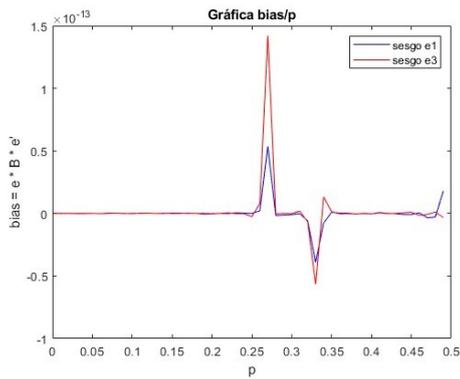
Bibliografía

- [1] Kuusela, Mikael. *Uncertainty quantification in unfolding elementary particle spectra at the Large Hadron Collider*. Escuela Politécnica Federal de Lausana, Lausana, Switzerland, July 23, 2016.
- [2] Kuusela, Mikael. *Unfolding: A Statistician's Perspective*. NuXTract 2023, CERN, Geneva, Switzerland, October 3, 2023.
- [3] A. N. Tikhonov, *Solution of incorrectly formulated problems and the regularization method*, Soviet Mathematics: Doklady, vol. 4, pp. 1035–1038, 1963.
- [4] Y. Vardi, L. A. Shepp, and L. Kaufman. *A statistical model for positron emission tomography*. Journal of the American Statistical Association, 80(389):8–20, 1985.
- [5] A. Höcker and V. Kartvelishvili. *SVD approach to data unfolding*. Nuclear Instruments and Methods in Physics Research A, 372:469–481, 1996.
- [6] S. Schmitt. *TUnfold, an algorithm for correcting migration effects in high energy physics*. Journal of Instrumentation, 7:T10003, 2012.
- [7] P. Dempster, N. M. Laird, and D. B. Rubin. *Maximum likelihood from incomplete data via the EM algorithm*. Journal of the Royal Statistical Society. Series B (Methodological), 39(1):1–38, 1977.
- [8] G. Cowan, *A Survey of Unfolding Methods for Particle Physics*, Physics Department, Royal Holloway, University of London, UK, Proceedings of Conference XYZ, pp. 248-253.
- [9] Kuusela, Mikael. *Statistical Methods for Unfolding*, Report A, Department of Mathematics and Statistics, University of Helsinki, Finland, 2011.
- [10] G. D'Agostini. *Improved Iterative Bayesian Unfolding*. Università “La Sapienza” and INFN, Rome, Italy, arXiv:1010.0632 [physics.data-an], 2010. Disponible en: <http://arxiv.org/abs/1010.0632v1>
- [11] M. Spano. *Lectures on Unfolding*. Grenoble, May 31st, 2012. Disponible en: https://example.com/link_to_document
- [12] García-Lomas, M. 2024. *Programas numéricos para Unfolding*. GitHub. <https://github.com/maximogarcialomas/Programas-num-ricos-para-Unfolding>

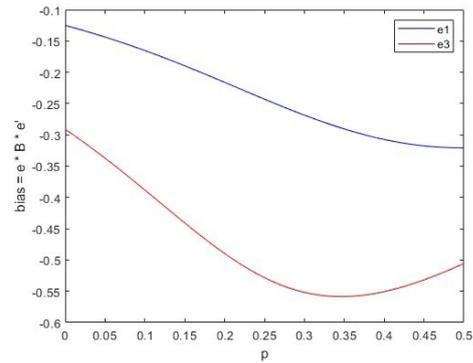
Apéndice A

Anexo A: Regularización de Tikhonov. Gráficas adicionales.

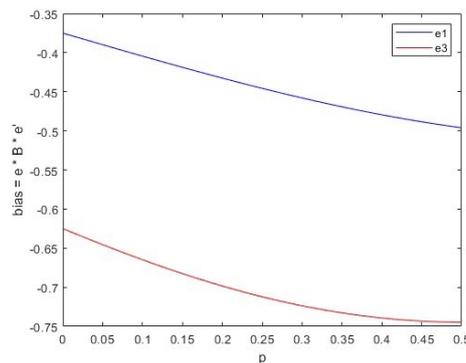
Gráficas adicionales para la regularización de Tikhonov que se han realizado para llegar a los resultados obtenidos en el cuerpo del estudio pero que no se han considerado relevantes suficientes para comentarlas en el cuerpo.



(a) $\delta = 0$

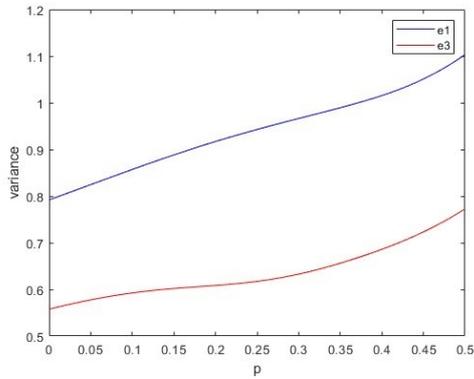


(b) $\delta = 0,1$

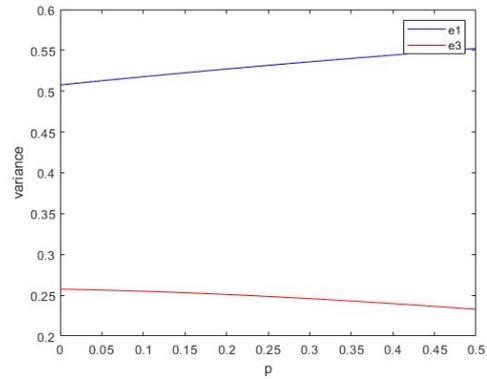


(c) $\delta = 1$

Figura A.1: Sesgo del canal externo e_1 y el central e_3 para $p \in [0, 0,5]$ en el caso 6×6 para el método de Tikhonov regularizado con la derivada para $\delta = 0, 0,1, 1$

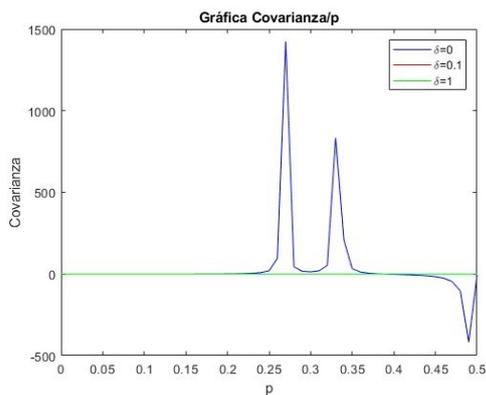


(a) $\delta = 0,1$

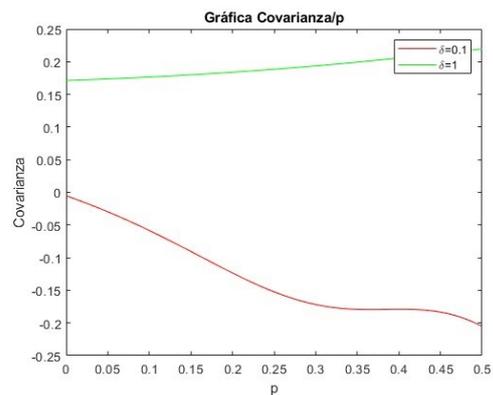


(b) $\delta = 1$

Figura A.2: Varianza del canal externo e_1 y del central e_3 para $p \in [0, 0,5]$ en el caso 6×6 para el método de Tikhonov regularizado con la derivada con $\delta = 0,1, 1$.



(a) $\delta = 0, 0,1, 1$



(b) $\delta = 0,1, 1$

Figura A.3: Covarianza entre el canal externo e_1 y del central e_3 para $p \in [0, 0,5]$ en el caso 6×6 para el método de Tikhonov regularizado con la derivada con $\delta = 0,1, 1$.

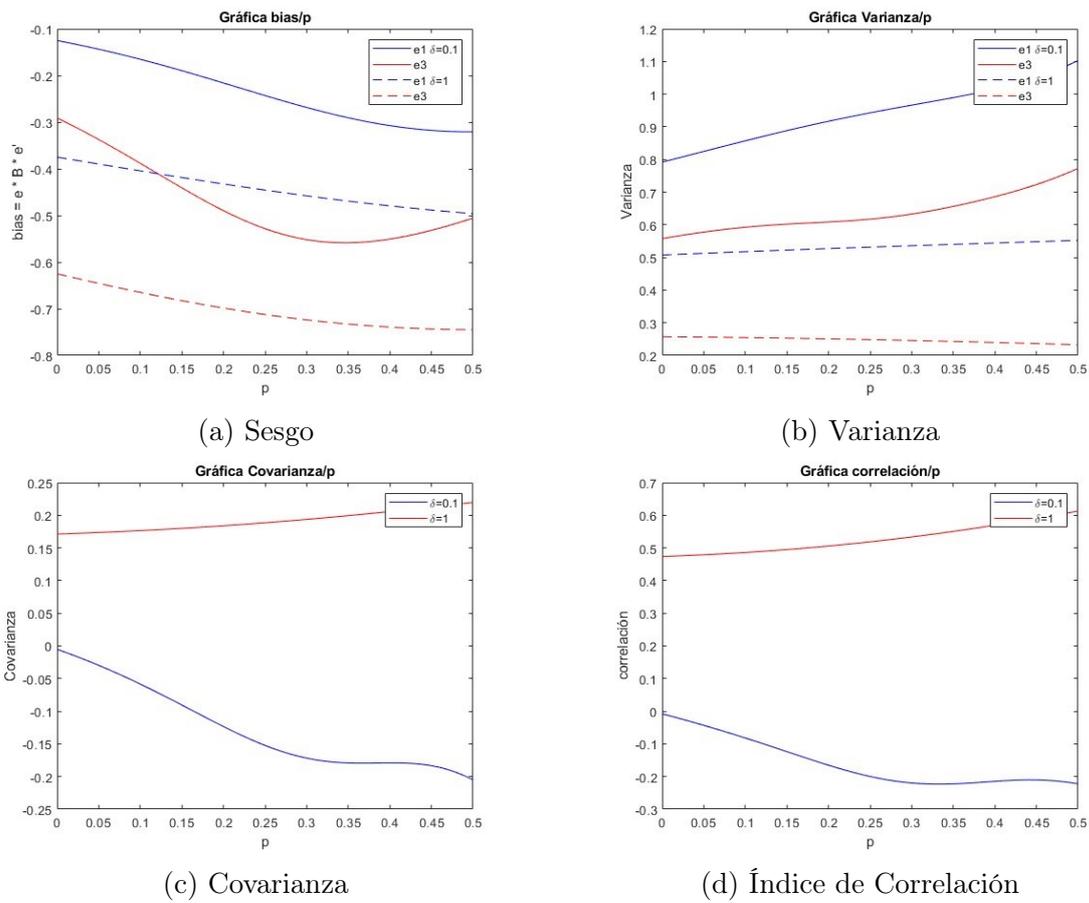


Figura A.4: Sesgo y varianza del canal externo e_1 y del central e_3 , así como la covarianza y la correlación entre ellos para $p \in [0, 0,5]$ en el caso 6×6 para el método de Tikhonov regularizado con SVD con $\lambda_{MC} = [1, 1, 1, 1, 1, 1]$ para $\delta = 0,1, 1$.

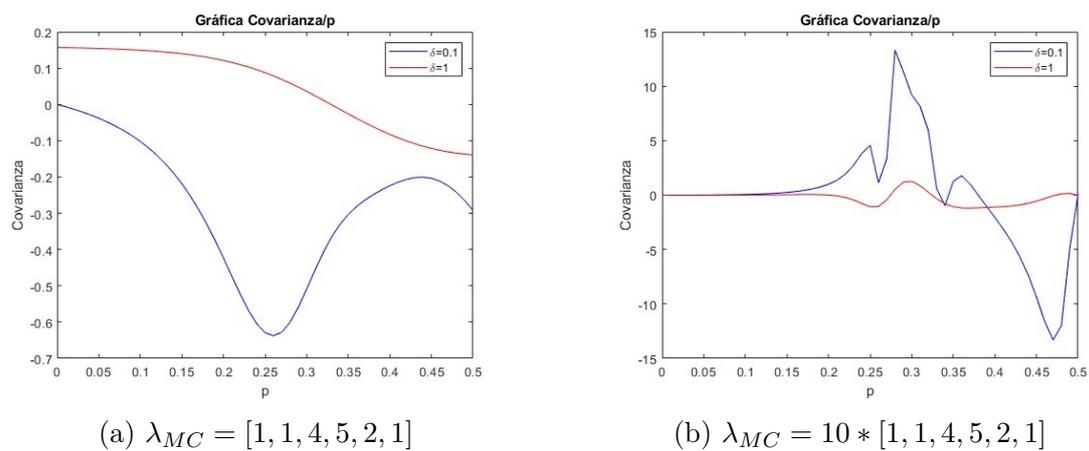


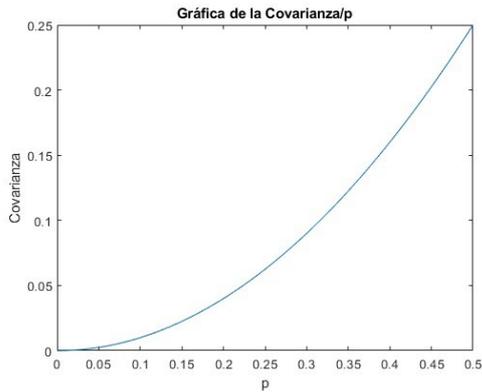
Figura A.5: Covarianza entre el canal externo e_1 y del central e_3 para $p \in [0, 0,5]$ en el caso 6×6 para el método de Tikhonov regularizado con SVD con $\lambda_{MC} = [1, 1, 4, 5, 2, 1]$ y $\lambda_{MC} = 10 * [1, 1, 4, 5, 2, 1]$ para $\delta = 0,1, 1$.

Apéndice B

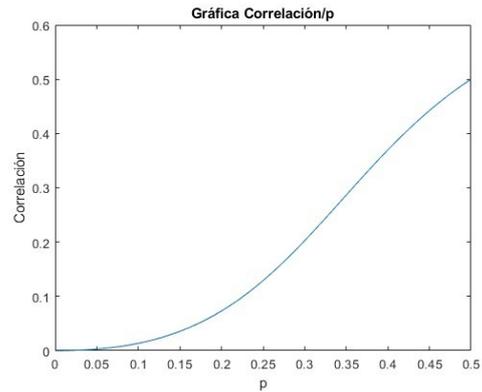
Anexo B: Método iterativo con 'early stopping'. Gráficas adicionales.

Gráficas adicionales para el método iterativo con 'early stopping' que se han realizado para llegar a los resultados obtenidos en el cuerpo del estudio pero que no se han considerado relevantes suficientes para comentarlas en el cuerpo.

$$\lambda_0 = [1, 1, 1, 1, 1, 1]$$

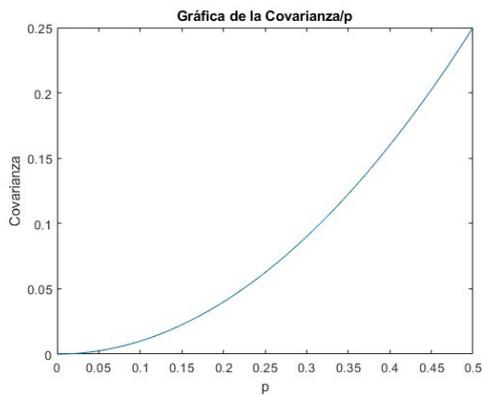


(a) Covarianza

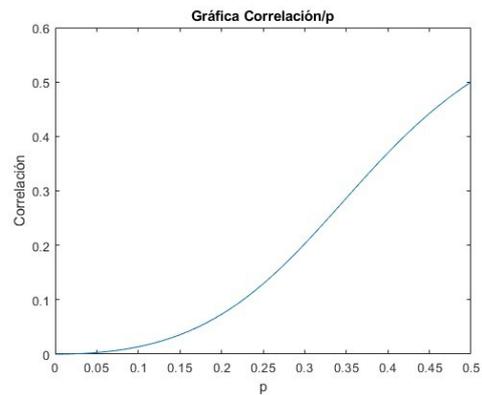


(b) Índice de Correlación

$$\lambda_0 = [1, 1, 4, 5, 2, 1]$$

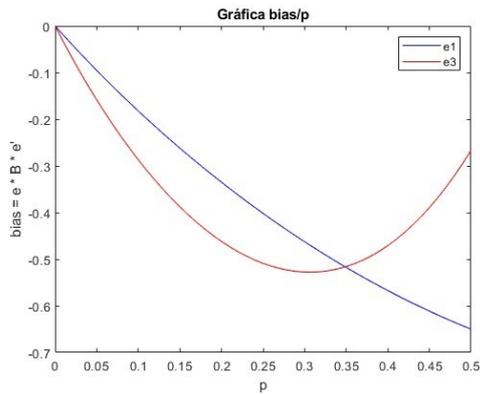


(c) Covarianza

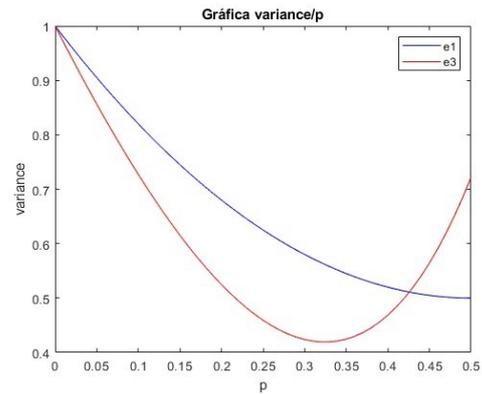


(d) Índice de Correlación

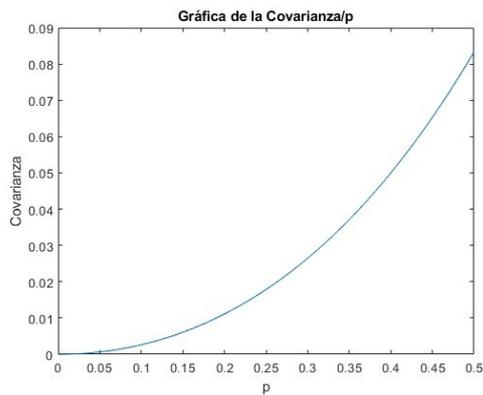
Figura B.1: Covarianza y la correlación entre el canal externo e_1 y el central e_3 para $p \in [0, 0,5]$ en el caso 6×6 para una iteración del método iterativo con 'early stopping' para $\lambda_0 = [1, 1, 1, 1, 1, 1]$ y $\lambda_0 = [1, 1, 4, 5, 2, 1]$



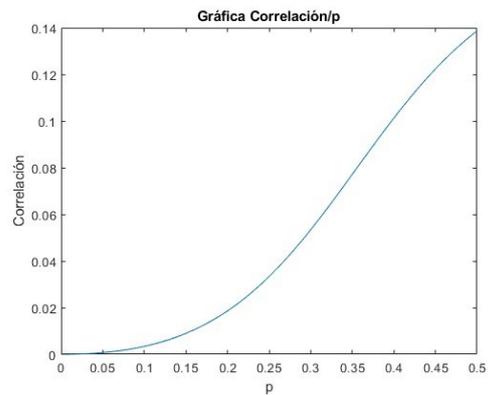
(a) Sesgo



(b) Varianza

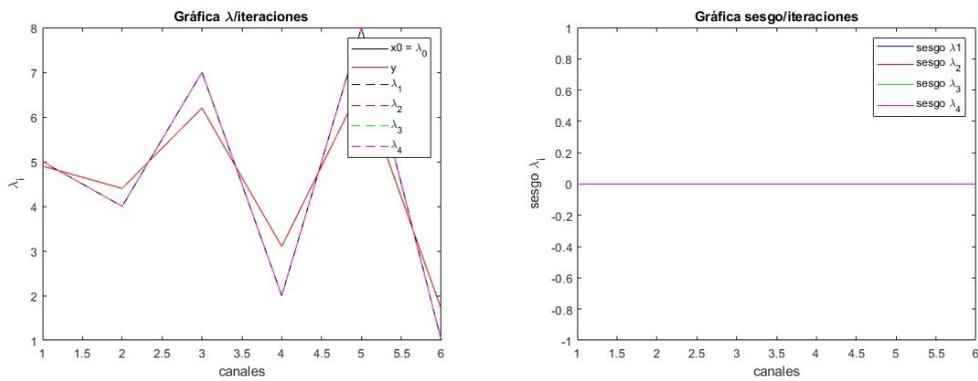


(c) Covarianza



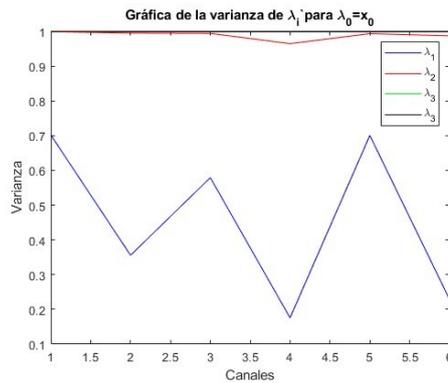
(d) Índice de Correlación

Figura B.2: Sesgo y varianza del canal externo e_1 y del central e_3 , así como la covarianza y la correlación entre ellos para $p \in [0, 0,5]$ en el caso 6×6 para una iteración del método iterativo con 'early stopping' para $\lambda_0 = 10 * [1, 1, 4, 5, 2, 1]$.



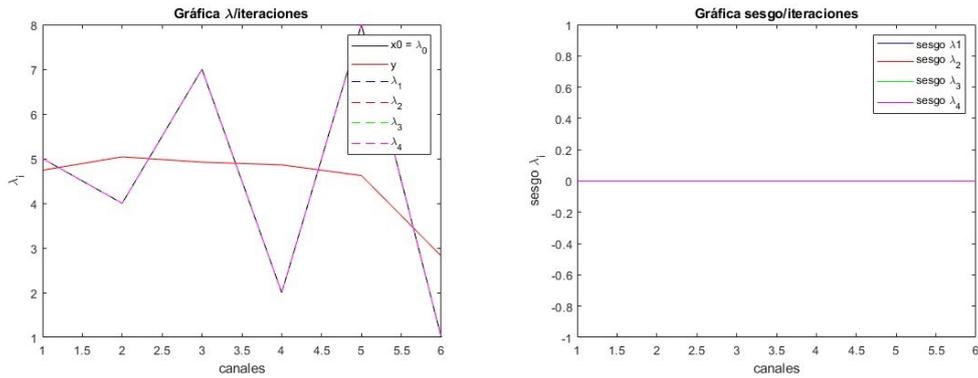
(a) Valores de λ_i

(b) Sesgo



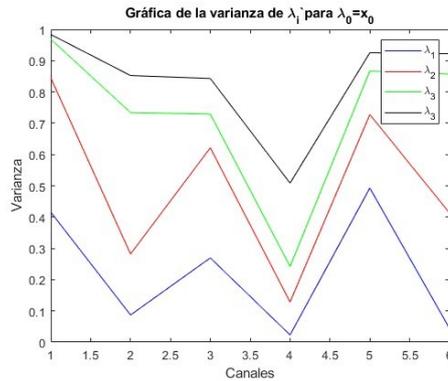
(c) Varianza

Figura B.3: Valores, sesgo y varianza de λ_i para 1, 10, 100, 1000 iteraciones del método iterativo con 'early stopping' con $p = 0,1$ partiendo de $\lambda_0 = x_0 = [5, 4, 7, 2, 8, 1]$



(a) Valores de λ_i

(b) Sesgo



(c) Varianza

Figura B.4: Valores, sesgo y varianza de λ_i para 1, 10, 100, 1000 iteraciones del método iterativo con 'early stopping' con $p = 0,26$ partiendo de $\lambda_0 = x_0 = [5, 4, 7, 2, 8, 1]$

Apéndice C

Anexo C: Intercambio sesgo-varianza (*Bias-Variance Trade Off*). Gráficas adicionales.

Gráficas adicionales para la función de intercambio sesgo-varianza que se han realizado para llegar a los resultados obtenidos en el cuerpo del estudio pero que no se han considerado relevantes suficientes para comentarlas en el cuerpo.

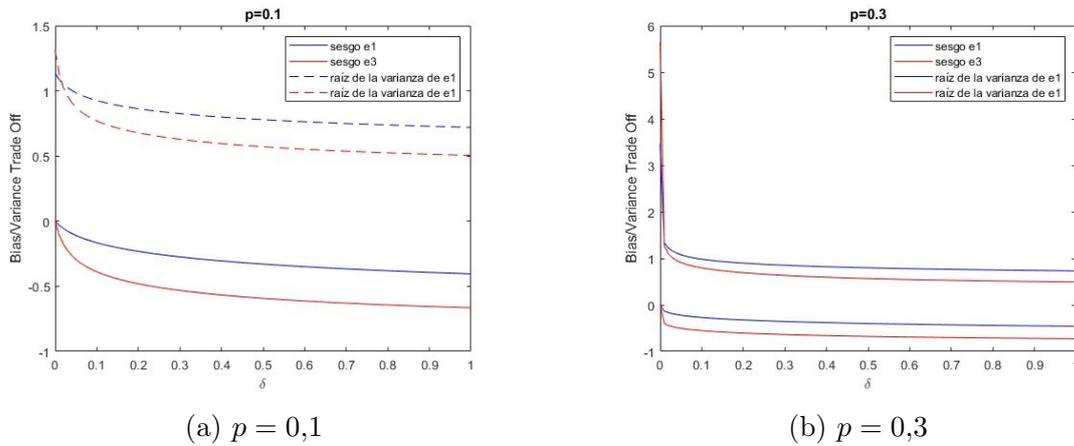
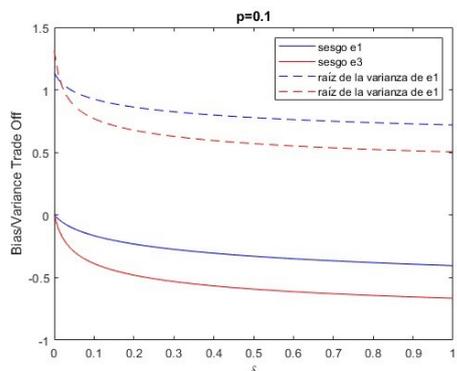
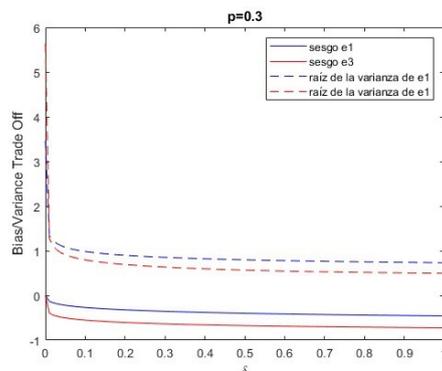


Figura C.1: Sesgo y la raíz de varianza del canal externo e_1 y del central e_3 para $p = 0,1, 0,3$ en el caso 6×6 para el método de Tikhonov regularizado con la derivada para $\delta \in [0, 1]$.

$$\lambda_{MC} = [1, 1, 1, 1, 1, 1]$$

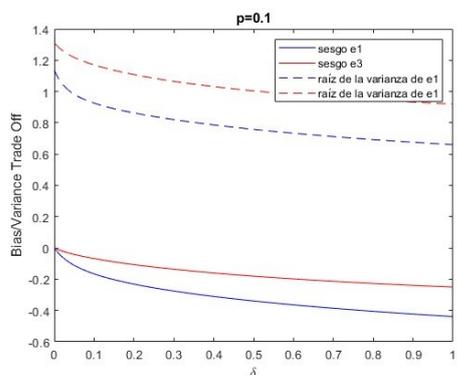


(a) $p = 0,1$

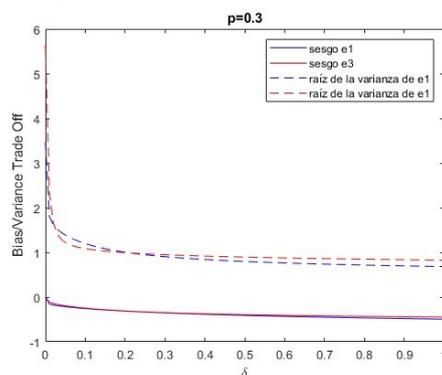


(b) $p = 0,3$

$$\lambda_{MC} = [1, 1, 4, 5, 2, 1]$$

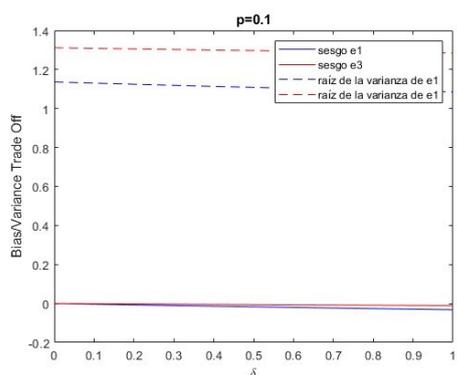


(c) $p = 0,1$

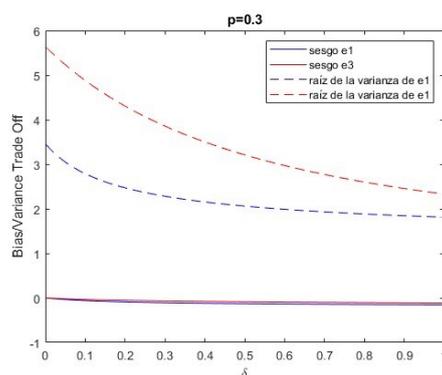


(d) $p = 0,3$

$$\lambda_{MC} = 10 * [1, 1, 4, 5, 2, 1]$$



(e) $p = 0,1$



(f) $p = 0,3$

Figura C.2: Sesgo y raíz de la varianza del canal externo e_1 y del central e_3 para $p = 0,1, 0,3$ en el caso 6×6 para el método de Tikhonov regularizado con SVD para $\delta \in [0, 1]$ y $\lambda_{MC} = [1, 1, 1, 1, 1, 1], [1, 1, 4, 5, 2, 1], 10 * [1, 1, 4, 5, 2, 1]$

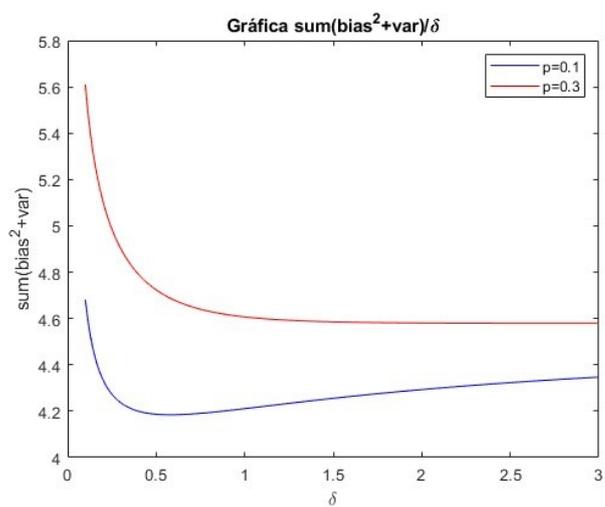


Figura C.3: Suma del cuadrado del sesgo más la varianza para $\delta \in [0,1,3]$, $p = 0,1$ y $p = 0,3$ en el caso 6×6 con el método SVD para $\lambda_{MC} = [1, 1, 1, 1, 1, 1]$.