Regularización del modelo para problemas inversos en exploración sísmica

Model regularization for linear inverse problems in exploration geophysics

Juan I Sabbione[1]

Resumen La sísmica de exploración permite caracterizar el subsuelo utilizando mediciones usualmente adquiridas en superficie. Las propiedades físicas del medio se pueden inferir a partir de los efectos sobre los datos observados. Tal estrategia en ciencia se conoce como resolución del problema inverso. El problema inverso se puede plantear como un problema de optimización en el que se determinan los parámetros que minimizan cierta función de costo. Muchos de los problemas inversos que se presentan en sísmica de exploración o bien son lineales, o bien pueden ser estudiados mediante una aproximación lineal. Para que un problema inverso esté bien definido, deben cumplirse tres condiciones: que la solución exista, que sea única, y que cambie en forma continua con los parámetros del modelo (i.e., que sea estable). Desafortunadamente, estas condiciones no se cumplen simultáneamente en ningún caso. En general, las soluciones buscadas no son únicas y/o son inestables. Esto conlleva a que el modelo a invertir deba ser regularizado. La regularización puede ser explícita o implícita. La primera se logra introduciendo un término de penalización sobre la función de costo. La segunda se puede alcanzar por detención temprana de los algoritmos iterativos, descartando *outliers*, o bien utilizando algoritmos voraces (*greedy*) más modernos. Aquí se presentan ejemplos de estas últimas técnicas a partir de la transformada de Radon.

Palabras clave Métodos sísmicos, problema inverso, regularización, algoritmos voraces, transformada de Radon.

Abstract *Exploration geophysics allows us to characterize the subsurface using measurements typically acquired at the surface. The physical properties of the medium can be inferred from the effects observed in the data. This approach in science is known as solving an inverse problem. The inverse problem can be formulated as an optimization problem, where parameters are determined to minimize a specific cost function. Many inverse problems in exploration geophysics are either linear or can be studied through a linear approximation. For an inverse problem to be well-posed, three conditions must be met: the solution must exist, it must be unique, and it should vary continuously with the model parameters (i.e., it should be stable). Unfortunately, these conditions are never all satisfied simultaneously. Generally, the solutions sought are either non-unique and/or unstable, which means the model to be inverted must be regularized. Regularization can be explicit or implicit. The former is achieved by introducing a penalty term in the cost function. The latter can be achieved by early stopping of iterative algorithms, discarding outliers, or using more modern greedy algorithms. Examples of these latter techniques are presented here using the Radon transform.*

Keywords *Seismic methods, inverse problem, regularization, greedy algorithms, Radon transform.*

INTRODUCCIÓN

Para poder inferir las propiedades del medio a partir de observaciones en superficie, primero se deben establecer las leyes físicas que gobiernan la propagación de ondas acústicas y elásticas a través del

[1]Facultad de Ciencias Astronómicas y Geofísicas, Universidad Nacional de La Plata, y CONICET, La Plata, Argentina. Email: jsabbione@fcaglp.unlp.edu.ar

subsuelo. Esto se denomina problema directo o modelado. Es decir, se buscan los efectos que se producen en la propagación de ondas cuando las causas son conocidas. Por otro lado, el enfoque del problema inverso es el opuesto: hallar los parámetros del modelo que generaron las observaciones. En otras palabras, el problema inverso se basa en determinar las causas que produjeron los efectos observados. Estos conceptos se pueden describir mediante el siguiente diagrama:

$$x \longrightarrow \boxed{G(m)} \longrightarrow y = G(m; x),$$

donde x representa la entrada, G el modelo que depende de los parámetros m, e y las observaciones o datos.

En un problema directo, conocido el modelo G y los parámetros que lo determinan m, se buscan derivar las ecuaciones para conocer la respuesta y del sistema ante cierta perturbación x. Esto es, dados m y x, hallar $y = G(m; x)$. Por el contrario, en un problema inverso se parte de la perturbación introducida x y de las observaciones y para hallar los parámetros m que caracterizan al modelo capaz de generar tales observaciones.

En el problema inverso, no siempre se conoce la perturbación de entrada x con exactitud. También puede ocurrir que tal perturbación esté contemplada dentro del modelo G. Por ello, es conveniente describir al problema de forma simplificada como

$$y = G(m). \tag{1}$$

Algunos ejemplos de problemas directos en exploración sísmica son: teoría de propagación de ondas, estudios de dispersión en física de rocas, propagación en medios porosos saturados (teoría de Biot), modelado sismoeléctrico, desarrollo de operadores de propagación en migración sísmica, etc. Asimismo, entre los ejemplos de problemas inversos, se pueden mencionar: estimación de ondícula, filtros inversos Wiener, transformadas de Radon, estimación de parámetros de anisotropía, construcción de imágenes mediante migración sísmica, e inversión de onda completa, entre otros.

EL PROBLEMA INVERSO LINEAL

En muchos casos, la dependencia del modelo G con los parámetros m o bien es lineal, o bien puede ser aproximada linealmente. Así, la ecuación 1 puede escribirse como:

$$d = L\, m, \tag{2}$$

donde L es el operador lineal que representa al modelo y d son los datos observados. El operador L se denomina operador directo, ya que genera datos d a partir del modelo. Análogamente, el operador transpuesto de L, que se denota como L^T, genera los parámetros o coeficientes adjuntos del modelo (m_a) a partir de los datos:

$$m_a \approx L^T\, d\,. \tag{3}$$

En las transformaciones auto-adjuntas, como es el caso de la transformada discreta de Fourier, el operador L^T es igual a su inverso: $L^T = L^{-1}$. Sin embargo, en la gran mayoría de las transformaciones esto no ocurre y $L^T \neq L^{-1}$.

La transformada de Radon (RT), que representa una herramienta muy flexible de procesamiento sísmico (Sabbione & Sacchi, 2016a), es un caso en el que puede verse claramente que la transformación adjunta no es la equivalente de la inversa. Por ello, para hallar m a partir de los datos se busca que los datos u observaciones d se parezcan lo más posible a aquello que predice el modelo: $L\, m$. Esto induce la necesidad de introducir un función de costo a minimizar y un métrica en el espacio de los datos. Dicha métrica queda definida por una norma matemática, y la selección de la norma es crucial para resolver el problema. Suponiendo que los errores se distribuyen normalmente, se suele utilizar la norma L2 y el problema inverso se convierte en un problema de cuadrados mínimos cuya solución es conocida:

$$\tilde{m} = \underset{m}{\mathrm{argmin}}\{\|Lm - d\|_2^2\} \quad \Rightarrow \quad \tilde{m} = (L^\intercal L)^{-1} L^\intercal d. \tag{4}$$

Desde el punto de vista del álgebra lineal, el operador L se puede explicitar por medio la matriz que representa la transformación lineal correspondiente. Sin embargo, existen muchos problemas para los que la matriz es demasiado grande y el operador lineal L debe entenderse como un programa de computación que ejecuta la acción de L sobre m, sin que nunca se forme la matriz.

REGULARIZACIÓN

Para que el problema dado por la ecuación 4 esté bien definido, deben cumplirse tres condiciones: (1) que exista la solución, (2) que la solución sea única, y (3) que la solución varíe en forma continua con las condiciones iniciales (condición de estabilidad). En la gran mayoría de los casos, alguna de la condiciones anteriores no se cumple. En general la solución existe, pero suele no ser única, y/o suele ser inestable. Esto induce la necesidad de *regularizar* el problema.

La solución usualmente adoptada es la regularización clásica de Tikhonov, también denominada de cuadrados mínimos amortiguados (DLS, *damped least-squares*). Esta solución está basada en agregar un término de regularización sobre la función de costo que penaliza la norma del modelo y lo convierte en un problema estable:

$$\tilde{m} = \underset{m}{\mathrm{argmin}}\{J(m) = \|Lm - d\|_2^2 + \mu\,\|m\|_2^2\} \quad \Rightarrow \quad \tilde{m} = (L^\intercal L + \mu\mathcal{I})^{-1} L^\intercal d\,, \tag{5}$$

donde $\mathcal{I}$ representa la identidad. Un ejemplo clásico de procesamiento sísmico en el que se suele adoptar esta solución es en el filtro Wiener (Treitel & Lines, 1982).

Regularización explícita

El tipo de regularización presentada en la ecuación 5 se denomina regularización explícita, ya que se logra agregando el término $\mu\|m\|_2^2$ en la función de costo. Se puede elegir la misma norma que en el término de error de ajuste $\|Lm - d\|_2^2$, como en la regularización de Tikhonov, o bien una distinta. Esto dependerá de cuál sea la característica que se desea imprimir sobre la solución buscada. En ocasiones, el modelo buscado debe ser ralo (*sparse*); es decir, con pocos elementos no nulos. Para ello, puede utilizarse una norma L1 o de Cauchy, que es la idea adoptada en la transformada de Radon de alta resolución de Sacchi & Ulrych (1995), y que se basa en proponer que bastan unos pocos coeficientes para representar las reflexiones. A diferencia de lo que ocurre en la regularización de Tikhonov, donde la solución se obtiene a partir de una fórmula cerrada (ecuación 5), cuando se elije otra norma para el término de regularización la solución debe encontrase a partir de algoritmos iterativos, como por ejemplo el de gradiente conjugado.

Regularización implícita

Cualquier otra forma de regularización que no involucre agregar un término extra sobre la función de costo penalizando cierta norma del modelo se denomina regularización implícita. Se puede optar por métodos que truncan un algoritmo iterativo antes de que converja para evitar ajustar errores, construir el modelo mediante pocos coeficientes (Pati et al., 1993), o métodos de tipo de aprendizaje automático (*machine learning*), entre otros. Estas técnicas resultan muy útiles en problemas de procesamiento sísmico en los que el modelo que se busca invertir suele ser muy denso y de varias dimensiones.

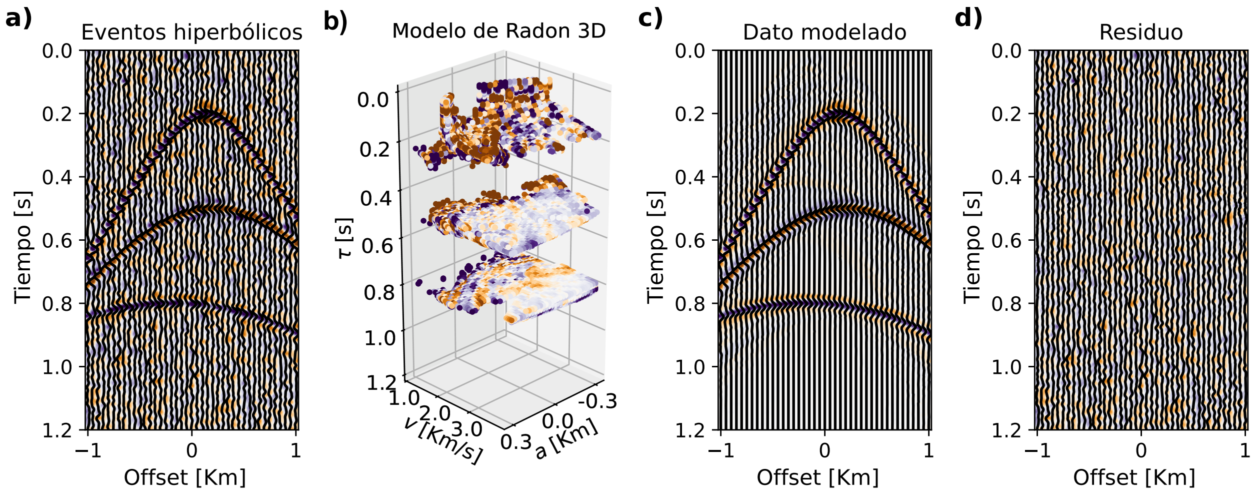

Figura 1. *Transformada de Radon hiperbólica con ápice corrido. a) Hipérbolas de reflexión con ruido. b) Modelo de Radon 3D invertido con dominio restringido. c) Datos modelados (atenuación de ruido). d) Residuo (datos de entrada − datos modelados).*

EJEMPLOS

La transformada de Radon en el dominio del tiempo resulta muy costosa computacionalmente. Por ello, es conveniente regularizar el modelo restringiendo su dominio para acelerar el procesamiento. Esta idea se basa en predecir los datos utilizando un pequeño subespacio del modelo a invertir. Tal estrategia permite además obtener muy buenos resultados de atenuación de ruido (*denoising*) (Sabbione & Sacchi, 2016b). Otra aplicación que utiliza este tipo de enfoque es la atenuación de ruido por transformada de Radon en datos microsísmicos. En microsísmica, el dominio de Radon es 4D para arreglos lineales dispuestos en pozos, y 5D para arreglos de receptores de superficie. Restringir el dominio del modelo regulariza el problema y acelera considerablemente la inversión (Sabbione et al., 2013, 2015).

En el campo de la sísmica convencional, la RT hiperbólica con ápice corrido representa otro ejemplo de este tipo de métodos, y puede ser utilizada para procesar datos de sísmica de reflexión. En este problema, los pares de operadores directo y transpuesto están dados por:

$$d(t,x) = \sum_{v,a} m\left(\tau = \sqrt{t^2 - \frac{(x-a)^2}{v^2}},\, v,\, a\right), \tag{6}$$

$$m_a(\tau,v,a) = \sum_{x} d\left(t = \sqrt{\tau^2 + \frac{(x-a)^2}{v^2}},\, x\right). \tag{7}$$

El modelo de Radon es 3D y está determinado por el tiempo de ida y vuelta τ, la velocidad efectiva v, y el ápice a. La Figura 1 demuestra que se pueden obtener excelentes resultados modelando los datos utilizando solo un subconjunto del dominio del modelo.

Basados mayormente en la idea de Orthogonal Matching Pursuit (OMP) (Pati et al., 1993), existen también diferentes métodos que regularizan el problema inverso a partir de construir el modelo iterativamente utilizando la menor cantidad posible de coeficientes a partir de cierto criterio. Estos métodos se denominan voraces (*greedy*). Entre ellos, seleccionamos para hacer pruebas de inversión del modelo los métodos de RT voraz (GRT) (Wang et al., 2010), RT vía orthogonal matching pursuit por etapas (StOMP) (Donoho et al., 2012), RT vía matching pursuit ortgonal regularizado (ROMP) (Needell & Vershynin, 2010), y RT vía búsqueda de subespacio (SP) (Dai & Milenkovic, 2009).

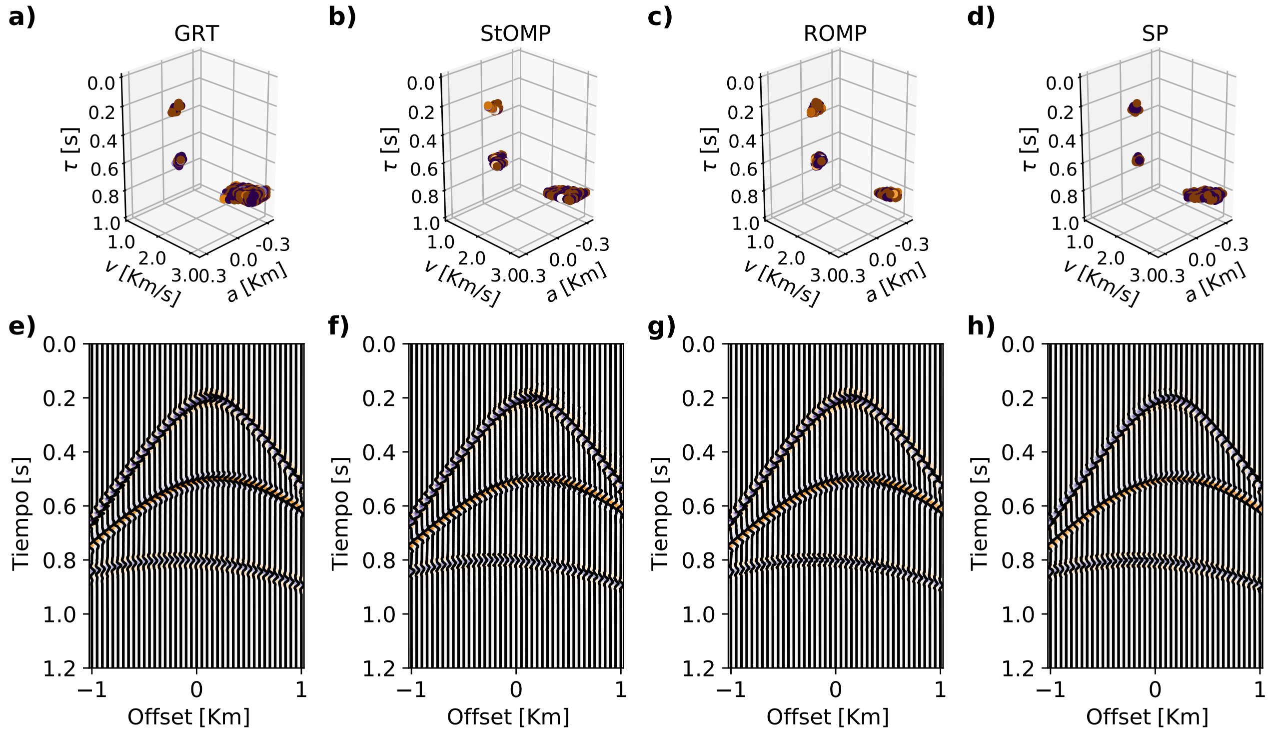

Figura 2. *Resultados de denoising mediante regularización por algoritmos voraces. Fila superior: modelo invertido mediante a) GRT, b) StOMP, c) ROMP y d) SP. Fila inferior: datos modelados vía e) GRT, f) StOMP, g) ROMP y h) SP.*

En la Figura 2 puede observarse cómo los algoritmos voraces utilizados para el cálculo de la RT en el mismo ejemplo mostrado en la Figura 1 permiten predecir muy bien los datos atenuando el ruido. Es notable que estos resultados se logran mediante muy pocos coeficientes en el modelo, lo que regulariza el problema de forma muy eficiente. Estas técnicas tienen aplicación en distintos problemas reales, como por ejemplo en la atenuación de múltiples en sísmica marina (Sabbione & Sacchi, 2017).

CONCLUSIONES

Una gran cantidad de problemas de exploración sísmica se pueden plantear como un problema inverso. Para su resolución, primero debe estudiarse el problema o modelado directo: analizar las observaciones o datos que genera un dado modelo. En otras palabras, cómo son los efectos a partir de las causas. Luego se puede abordar el problema inverso: conocer los parámetros del modelo a partir de los datos u observaciones. Esto es, encontrar cuáles serían las causas que producen los efectos observados. En muchos problemas inversos el modelo se comporta linealmente respecto de los parámetros y puede ser planteado como un problema de optimización lineal. Para resolverlo, se deben definir dos operadores: un operador directo, que genera los datos u observaciones a partir del modelo; y un operador adjunto, que estima los coeficientes o parámetros del modelo a partir de los datos. Muchos problemas inversos están mal condicionados y deben ser regularizados para que su solución sea única y estable.

La regularización puede hacerse de dos maneras: (1) Explícitamente, agregando un término de penalización sobre la función de costo con una norma adecuada; o (2) implícitamente, a partir de truncar las iteraciones de la inversión en forma temprana o bien construyendo el modelo iterativamente con pocos coeficientes. Muchas aplicaciones geofísicas se caracterizan porque el modelo a invertir es ralo (*sparse*). Esta cualidad del modelo puede imponerse a partir de agregar un término de regularización apropiado sobre la función de costo a minimizar.

Existen dos grandes grupos de algoritmos para lograr que el modelo sea *sparse*: los que utilizan norma de tipo L1 (o Cauchy), y aquellos que buscan construir el modelo iterativamente con la menor cantidad de elementos posible (denominados métodos voraces). Las diferentes estrategias de regularización del modelo se pueden utilizar en una amplia variedad de problemas geofísicos: filtros y otros problemas de procesamiento, atenuación de ruido en sísmica convencional y microsísmica, atenuación de múltiples en sísmica marina, migración, etc. En particular, los métodos voraces (*greedy*) ofrecen el valor agregado de acelerar la inversión ya que el modelo es invertido utilizando la menor cantidad posible de coeficientes según criterios determinados por el algoritmo voraz utilizado.

REFERENCIAS

Dai, W. & Milenkovic, O. (2009). Subspace pursuit for compressive sensing signal reconstruction. *IEEE Trans. Inf. Theor.*, 55(5), 2230–2249. https://doi.org/10.1109/TIT.2009.2016006

Donoho, D. L., Tsaig, Y., Drori, I., & Starck, J. L. (2012). Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit. *IEEE Trans. Inf. Theor.*, 58(2), 1094–1121. https://doi.org/10.1109/TIT.2011.2173241

Needell, D. & Vershynin, R. (2010). Signal recovery from incomplete and inaccurate measurements via regularized orthogonal matching pursuit. *IEEE Journal of selected topics in signal processing*, 4(2), 310–316. https://doi.org/10.1190/1.1443845

Pati, Y., Rezaiifar, R., & Krishnaprasad, P. (1993). Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition. *Proceedings of the 27th Annual Asilomar Conference on Signals, Systems, and Computers*, 40–44 vol. 1. https://doi.org/10.1109/ACSSC.1993.342465

Sabbione, J. I. & Sacchi, M. D. (2016a). Fast time domain hyperbolic Radon transforms. *GeoConvention 2016*, 1–5.

Sabbione, J. I. & Sacchi, M. D. (2016b). Restricted model domain time Radon transforms. *Geophysics*, 81(6), A17–A21. https://doi.org/10.1190/geo2016-0270.1

Sabbione, J. I. & Sacchi, M. D. (2017). Attenuating multiples with the restricted domain hyperbolic Radon transform. *15th International Congress of the Brazilian Geophysical Society*, 603–608. https://doi.org/10.1190/sbgf2017-118

Sabbione, J. I., Sacchi, M. D., & Velis, D. R. (2015). Radon transform–based microseismic event detection and signal–to–noise ratio enhancement. *Journal of Applied Geophysics*, 113, 51–63. https://doi.org/10.1016/j.jappgeo.2014.12.008

Sabbione, J. I., Velis, D. R., & Sacchi, M. D. (2013). Microseismic data denoising via an apex–shifted hyperbolic Radon transform. *SEG Technical Program Expanded Abstracts 2013*, 2155–2161. https://doi.org/10.1190/segam2013-1414.1

Sacchi, M. D. & Ulrych, T. J. (1995). High-resolution velocity gathers and offset space reconstruction. *Geophysics*, 60(4), 1169–1177. https://doi.org/10.1190/1.1443845

Treitel, S. & Lines, L. R. (1982). Linear inverse theory and deconvolution. *Geophysics*, 47(8), 1153–1159. https://doi.org/10.1190/1.1441378

Wang, J., Ng, M., & Perz, M. (2010). Seismic data interpolation by greedy local Radon transform. *Geophysics*, 75(6), WB225–WB234. https://doi.org/10.1190/1.3484195