Análisis de la convergencia del método VarPro para problemas inversos no lineales separables regularizados

Gabriela Jeronimo
Universidad de Buenos Aires & CONICET, Argentina
jeronimo@dm.uba.ar

Consideramos problemas inversos de la forma $\mathbf{A}(\mathbf{y})\mathbf{x} = \mathbf{b} = \mathbf{b}_{\text{true}} + \epsilon$ con $\mathbf{A}(\mathbf{y})\mathbf{x}_{\text{true}} = \mathbf{b}_{\text{true}}$, donde $\mathbf{b}_{\text{true}} \in \mathbb{R}^m$ denota un vector desconocido asociado a los datos y $\epsilon \in \mathbb{R}^m$ es un vector desconocido que representa ruido o errores. La matriz $\mathbf{A}(\mathbf{y}) \in \mathbb{R}^{m\times n}$ con $m \geq n$ es desconocida, pero suponemos que puede parametrizarse en forma no lineal por un vector $\mathbf{y} \in \mathbb{R}^r$ con $r \ll n$. A estos problemas los llamamos problemas inversos no lineales separables. Dado un vector de datos $\mathbf{b}$ y una función matricial $\mathbf{A}(\mathbf{y})$, el objetivo es calcular buenas aproximaciones de $\mathbf{x}$ y de $\mathbf{y}$. Para esto, se busca resolver $\min_{\mathbf{x},\mathbf{y}} \|\mathbf{A}(\mathbf{y})\mathbf{x} - \mathbf{b}\|_2^2$. Nos interesamos en problemas ill-posed, para lo cual planteamos una formulación regularizada:

$$\min_{\mathbf{x},\mathbf{y}} \|\mathbf{A}(\mathbf{y})\mathbf{x} - \mathbf{b}\|_2^2 + \lambda^2\|\mathbf{x}\|_2^2,$$

donde $\lambda > 0$ es un parámetro de regularización. Nos enfocamos en el método de proyección de variables (VarPro) introducido en [1]. La idea fundamental consiste en eliminar la variable $\mathbf{x}$ (resolviendo un problema lineal de cuadrados mínimos para cada $\mathbf{y}$), reemplazar $\mathbf{x} = \mathbf{x}(\mathbf{y})$, y finalmente resolver un problema de minimización $\min_{\mathbf{y}} \|\mathbf{F}(\mathbf{y})\|^2$ sólo en las variables $\mathbf{y}$. Este problema no lineal puede resolverse por el método de Gauss-Newton, lo que requiere calcular el Jacobiano de $\mathbf{F}$. Como esto suele ser muy costoso, se han utilizado distintas aproximaciones de este Jacobiano que mostraron buenos resultados computacionales, entre las que se destacan las propuestas por Kaufman ([2]) y por Ruano, Jones y Flemming ([3]).

En esta comunicación describiremos y analizaremos la convergencia de una nueva generalización del método VarPro aplicable a problemas de gran tamaño. Esta generalización consiste en calcular, para cada $\mathbf{y}$, una aproximación suficientemente buena de la solución $\mathbf{x} = \mathbf{x}(\mathbf{y})$ del problema de minimización correspondiente por medio de un método iterativo. Presentaremos un análisis teórico de la convergencia del método utilizando las aproximaciones de los Jacobianos de [2] y [3] calculadas, a su vez, con soluciones $\mathbf{x}$ aproximadas. Finalmente, comentaremos sobre experimentos numéricos cuyos resultados reflejan nuestro análisis teórico.

Trabajo en conjunto con Malena I. Español (Arizona State University, Estados Unidos).

Referencias

[1] G.H. Golub, V. Pereyra, The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate. SIAM J. Numer. Anal. 10 (1973), 413-432.

[2] L. Kaufman, A variable projection method for solving separable nonlinear least squares problems, BIT 15 (1975) 49-57.

[3] A.E.B. Ruano, D.I. Jones, P.J. Fleming, A new formulation of the learning problem of a neural network controller, in Proceedings of the 30th IEEE Conference on Decision and Control, 1991, 865-866.