

Conformación ciega de haz mediante regresión con máquinas de vectores soporte

Javier Merino, Ignacio Santamaría
Departamento de Ingeniería de Comunicaciones
Universidad de Cantabria
e-mail: merino@gtas.dicom.unican.es

Abstract—Blind beamforming is a common problem in wireless communications, where an array of antennas receives a number of signals from distinct locations at the same frequency and at the same time. In this paper the problem of blind beamforming for multiple constant modulus (CM) signals separation is solved using support vector machine (SVM) techniques. The CM property of the signal is used to formulate a regression problem which can be adapted to the SVM scheme, leading to an iterative reweighted algorithm. Once a signal is recovered, its contribution to the original observations is removed and the iterative procedure can be applied again to extract another CM signal. Simulation results show that this SVM-based algorithm offers better performance than the algebraic constant modulus algorithm (ACMA), mainly when only a small number of snapshots is available.

I. INTRODUCCIÓN

El problema de conformación de haz se presenta cuando varias señales son emitidas simultáneamente desde distintos lugares utilizando la misma frecuencia. Las señales son captadas por un conjunto de antenas; el objetivo del conformador de haz es aprovechar la diversidad espacial combinando linealmente las señales recibidas por las distintas antenas para extraer las señales emitidas originalmente [1].

Cuando no es posible disponer de una secuencia de entrenamiento para resolver el problema, es necesario aplicar algoritmos de conformación *ciega* de haz. Un caso particular de especial interés es la separación espacial de señales módulo constante (por ejemplo, QPSK). En este caso, con el objetivo de restaurar la propiedad de módulo constante a la salida del conformador, es posible aplicar el “Constant Modulus Algorithm” (CMA) en el entrenamiento de los distintos conformadores [2]. Se han propuesto distintas variantes e implementaciones (bloque e iterativas) del CMA en conformación de haz [3], [4], [5]. En especial, el “Analytical Constant Modulus Algorithm” (ACMA) [6] resuelve un problema bloque, encontrando la solución simultánea para P conformadores mediante la resolución de un problema de autovalores generalizado. El ACMA es un algoritmo robusto en presencia de ruido y proporciona resultados satisfactorios con un número reducido de observaciones (“snapshots”).

Recientemente, algoritmos basados en regresión mediante máquinas de vectores soporte (“Support Vector Machines”) se han propuesto para la igualación ciega de señales módulo constante [7], [8]. En este artículo se extienden las ideas allí propuestas al problema de la conformación ciega de haz de

señales módulo constante. Pese a su elevado coste computacional, esta técnica aprovecha las ventajas ofrecidas por las SVM en problemas de regresión requiriendo, en consecuencia, un conjunto de observaciones más reducido que el ACMA, como muestran las simulaciones realizadas.

II. FORMULACIÓN DEL PROBLEMA

Consideramos un conjunto de L señales que inciden simultáneamente sobre un array lineal compuesto por M antenas. Las observaciones a la salida del array se pueden modelar como

$$\mathbf{X} = \mathbf{A}\mathbf{S} + \mathbf{N}, \quad (1)$$

donde \mathbf{X} es una matriz de dimensiones $M \times N$ que contiene N muestras de las señales (banda base) captadas por cada una de las M antenas, \mathbf{S} es una matriz de dimensiones $L \times N$ con las señales de interés, y la matriz \mathbf{A} de dimensiones $M \times L$ representa la respuesta del array. La matriz $\mathbf{N} : M \times N$ corresponde al ruido aditivo presente en las observaciones, que modelamos como espacialmente blanco y gaussiano. Para un array lineal uniforme con antenas omnidireccionales, la matriz \mathbf{A} es de la forma

$$\mathbf{A} = \begin{bmatrix} 1 & \dots & 1 \\ e^{-j\phi_1} & \dots & e^{-j\phi_L} \\ \vdots & \vdots & \vdots \\ e^{-j(N-1)\phi_1} & \dots & e^{-j(N-1)\phi_L} \end{bmatrix}, \quad (2)$$

donde

$$\phi_i = 2\pi \frac{d}{\lambda} \sin \theta_i, \quad (3)$$

para $i = 1, 2, \dots, L$; siendo d la separación entre antenas, λ la longitud de onda de las señales y θ_i el ángulo de llegada de la señal i -ésima [1].

En este planteamiento suponemos que un número P (conocido) de entre las L señales incidentes tienen módulo constante. Nuestro problema consiste entonces en encontrar P conformadores \mathbf{w}_j , $j = 1, \dots, P$ que proporcionen las estimas de las señales originales,

$$y_j[n] = \sum_{i=1}^M w_{i,j} x_i[n] = \mathbf{w}_j^T \mathbf{x}[n], \quad (4)$$

para $j = 1, \dots, P$ y $n = 0, \dots, N - 1$, donde $\mathbf{x}[n]$ es la columna n -ésima de \mathbf{X} .

III. CONFORMACIÓN CIEGA MEDIANTE SVM

Para simplificar la notación consideramos en esta sección que únicamente una de las señales incidentes tiene módulo constante. El objetivo del conformador de haz es recuperar la propiedad de módulo constante de la señal digital de comunicaciones, esto es, $\|\mathbf{w}^T \mathbf{x}[n]\|^2 = 1$, para $n = 0, \dots, N-1$.

En el contexto de regresión mediante SVM, esto implica minimizar la función de coste

$$J(\mathbf{w}) = \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{n=0}^{N-1} |1 - \|\mathbf{w}^T \mathbf{x}[n]\|^2|_{\epsilon}, \quad (5)$$

donde

$$|1 - \|\mathbf{w}^T \mathbf{x}[n]\|^2|_{\epsilon} = \max\{0, |1 - \|\mathbf{w}^T \mathbf{x}[n]\|^2| - \epsilon\} \quad (6)$$

es la función de pérdidas de Vapnik, y C es una constante de regularización. Este problema es equivalente a minimizar la función

$$L(\mathbf{w}, \xi, \tilde{\xi}) = \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{n=0}^{N-1} (\xi_n + \tilde{\xi}_n), \quad (7)$$

con las restricciones

$$\|\mathbf{w}^T \mathbf{x}[n]\|^2 - 1 \leq \epsilon + \xi_n, \quad (8)$$

$$1 - \|\mathbf{w}^T \mathbf{x}[n]\|^2 \leq \epsilon + \tilde{\xi}_n, \quad (9)$$

$$\xi_n, \tilde{\xi}_n \geq 0, \quad (10)$$

para $n = 0, \dots, N-1$.

En la formulación convencional de regresión con SVM las restricciones son lineales respecto a las incógnitas \mathbf{w} , derivándose a partir de ahí un problema de programación cuadrática (QP) con una única solución y que puede ser resuelto eficientemente [10]. Por el contrario, las desigualdades (8) y (9) son cuadráticas respecto a los pesos del conformador. Para solucionar este problema empleamos el procedimiento propuesto en [7]. En particular, es posible reescribir el módulo cuadrado a la salida del conformador como

$$\|\mathbf{w}^T \mathbf{x}[n]\|^2 = \tilde{\mathbf{w}}^T \mathbf{g}[n], \quad (11)$$

donde se han definido los vectores

$$\tilde{\mathbf{w}} = \begin{bmatrix} \text{Re}(\mathbf{w}) \\ \text{Im}(\mathbf{w}) \end{bmatrix}, \quad \mathbf{g}[n] = \begin{bmatrix} \text{Re}(y[n] \mathbf{x}[n]^*) \\ \text{Im}(y[n] \mathbf{x}[n]^*) \end{bmatrix}, \quad (12)$$

donde $(\cdot)^*$ denota complejo conjugado. De esta manera, las desigualdades (8) y (9) se transforman en

$$\tilde{\mathbf{w}}^T \mathbf{g}[n] - 1 \leq \epsilon + \xi_n, \quad (13)$$

$$1 - \tilde{\mathbf{w}}^T \mathbf{g}[n] \leq \epsilon + \tilde{\xi}_n, \quad (14)$$

que son lineales respecto a $\tilde{\mathbf{w}}$ si se considera fija la salida del conformador $y[n]$. A partir de aquí el problema se puede formular como una regresión convencional con SVM. En concreto, se obtiene el siguiente problema QP: maximizar

$$W(\boldsymbol{\alpha}, \tilde{\boldsymbol{\alpha}}) = \sum_{n=0}^{N-1} (\tilde{\alpha}_n - \alpha_n) - \epsilon \sum_{n=0}^{N-1} (\alpha_n + \tilde{\alpha}_n) \quad (15)$$

$$- \frac{1}{2} \sum_{n,m=0}^{N-1} (\alpha_n - \tilde{\alpha}_n)(\alpha_m - \tilde{\alpha}_m) \langle \mathbf{g}[n], \mathbf{g}[m] \rangle,$$

sujeto a las restricciones $0 \leq \alpha_n, \tilde{\alpha}_n \leq C$; y siendo los patrones de entrada $\mathbf{g}[n]$ los definidos en (12). Se consigue así, por tanto, reformularlo como un problema QP convencional con variables reales cuya solución viene dada por

$$\tilde{\mathbf{w}} = \sum_{n=0}^{N-1} (\tilde{\alpha}_n - \alpha_n) \mathbf{g}[n]. \quad (16)$$

Finalmente, los coeficientes del conformador pueden expresarse como

$$\mathbf{w} = \sum_{n=0}^{N-1} \beta_n \mathbf{x}[n]^*, \quad (17)$$

donde se han definido unos multiplicadores de Lagrange ponderados como

$$\beta_n = (\tilde{\alpha}_n - \alpha_n) y[n]. \quad (18)$$

En general, el regresor lineal en el marco de las SVM incluye un término de umbral: $y[n] = \mathbf{w}^T \mathbf{x}[n] + b$. Para este problema, sin embargo, es necesario fijar $b = 0$, ya que de otra forma siempre se obtendría la solución trivial $\mathbf{w} = \mathbf{0}$ y $b = 1$.

IV. RESOLUCIÓN ITERATIVA

El problema formulado en la sección anterior no puede resolverse de forma directa en un sólo paso, ya que la salida del conformador es función de la propia solución,

$$y[n] = \mathbf{w}^T \mathbf{x}[n] = \sum_{k=0}^{N-1} (\tilde{\alpha}_k - \alpha_k) y[k] \mathbf{x}[k]^H \mathbf{x}[n], \quad (19)$$

donde $(\cdot)^H$ indica traspuesto conjugado. Es necesario, por lo tanto, aplicar un procedimiento iterativo como el presentado en [7]. Esta técnica se denomina ‘‘Iterative Reweighted Quadratic Programming’’ (IRWQP), debido a su similitud con el algoritmo ‘‘Iterative Reweighted Least Squares’’ (IRWLS) utilizado en algunos problemas de aproximación y regresión [9].

El método IRWQP se resume en tres pasos:

- 1) Resolver el problema QP (15) asumiendo fijo el valor de $y[n]$.
- 2) Obtener los nuevos coeficientes del conformador de haz según (17) y con ellos la nueva salida $y[n]$.
- 3) Repetir hasta converger.

El algoritmo se completa con un mecanismo de suavizado de los coeficientes del conformador entre iteración e iteración; de esta manera se evita la oscilación, que de otro modo tendría lugar, de los valores de la salida entre $y[n]$ y $1/y[n]$. En concreto, los coeficientes que se utilizan como solución del problema QP en la iteración k se obtienen como

$$\mathbf{w}_k = \lambda \mathbf{w}_{k-1} + (1 - \lambda) \mathbf{w}_{\text{QP}}, \quad (20)$$

donde \mathbf{w}_{QP} son los coeficientes resultantes de aplicar (17) a la solución del problema QP en la iteración k -ésima, y λ es el parámetro de suavizado.

V. EXTRACCIÓN DE MÚLTIPLES SEÑALES

El método descrito en la sección anterior permite extraer una de las señales de módulo constante que inciden en el array. En esta sección se propone un procedimiento para la extracción simultánea de P señales.

El algoritmo para la extracción simultánea se divide en dos etapas: una primera etapa de *inicialización* que apunta los P conformadores hacia cada una de las distintas señales. En la segunda fase del algoritmo cada conformador continúa iterando según el método convencional propuesto en la Sección IV, para refinar las soluciones obtenidas: es la etapa de *convergencia*.

La idea empleada para conseguir una inicialización adecuada consiste en eliminar de las observaciones de entrada al conformador k -ésimo, mediante un cancelador, la señal extraída por los conformadores anteriores $1, \dots, k-1$. De esta manera se evita que dos conformadores extraigan la misma señal. En particular, denotando la salida del conformador k -ésimo mediante el vector

$$\mathbf{y}_k = [y_k[0], y_k[1], \dots, y_k[N-1]]^T, \quad (21)$$

la matriz de datos que el conformador $k+1$ tomará como entrada durante la etapa de inicialización es

$$\mathbf{X}_{k+1} = \mathbf{X}_k - \mathbf{u}_k \mathbf{y}_k^T, \quad (22)$$

siendo $\mathbf{X}_1 = \mathbf{X}$ y donde

$$\mathbf{u}_k = [u_{1,k}, u_{2,k}, \dots, u_{M,k}]^T \quad (23)$$

es el vector de coeficientes complejos del cancelador de señal situado a la salida del conformador k -ésimo. El criterio para determinar los valores de estos coeficientes consiste en minimizar la norma L_2 de cada fila de la matriz \mathbf{X}_{k+1} , lo que conduce a

$$\mathbf{u}_k = \frac{\mathbf{X}_k \mathbf{y}_k^H}{\|\mathbf{y}_k\|_2^2}. \quad (24)$$

Tras un número fijo de iteraciones $niter1$ de la etapa de inicialización cada conformador apunta aproximadamente a cada una de las señales módulo constante incidentes. Entonces comienza la etapa de convergencia, en la que se prescinde del mecanismo de cancelación de señales y cada conformador aplica un número de iteraciones $niter2$ de forma independiente al resto.

Finalmente, el método propuesto para la separación simultánea de P señales de módulo constante se resume en el cuadro Algoritmo 1.

VI. RESULTADOS

En esta sección se compara el método propuesto con el ACMA [6]. Se ha considerado un ejemplo en el que cuatro señales, tres de ellas QPSK (módulo constante) y otra con distribución gaussiana, inciden sobre un array uniforme de $M = 4$ antenas omnidireccionales. Los ángulos de llegada de las señales son: 0° , 30° y 60° para las señales QPSK, y -20° para la de módulo no constante. La relación señal a ruido es de 15 dB para todas las señales. Se han considerado distintos

Algoritmo 1 Conformación ciega de haz con regresión SVM

Inicializar C , ϵ , λ y $\mathbf{w}_{0,k}$.

Etapa de inicialización

para $p = 1, 2, \dots, niter1$

para $k = 1, 2, \dots, P$

Calcular \mathbf{y}_k para $\mathbf{w}_{p-1,k}$ y \mathbf{X}_k .

Resolver el problema QP (15) y obtener \mathbf{w}_{QP} .

$\mathbf{w}_{p,k} = \lambda \mathbf{w}_{p-1,k} + (1 - \lambda) \mathbf{w}_{QP}$.

Calcular \mathbf{y}_k para $\mathbf{w}_{p,k}$, y \mathbf{X}_{k+1} mediante (22) y (24).

fin

fin

Etapa de convergencia

para $p = 1, 2, \dots, niter2$

para $k = 1, 2, \dots, P$

Calcular \mathbf{y}_k para $\mathbf{w}_{p-1,k}$ y \mathbf{X} .

Resolver el problema QP (15) y obtener \mathbf{w}_{QP} .

$\mathbf{w}_{p,k} = \lambda \mathbf{w}_{p-1,k} + (1 - \lambda) \mathbf{w}_{QP}$.

fin

fin

ejemplos variando el número de observaciones (“snapshots”) $N = 8, 10, 13, 16, 25, 50$. Para cada longitud se han promediado los resultados de 300 simulaciones independientes.

Para la técnica de conformación basada en SVM se han escogido los siguientes valores: $\epsilon = 0.01$ (parámetro de la función de Vapnik) y $\lambda = 0.3$ (factor de suavizado). El parámetro de regularización C , a su vez, se ha estimado de acuerdo a la expresión

$$C = \frac{1}{M} \sum_{m=1}^M (\overline{g_m[n]} + 3\sigma_{g_m}), \quad (25)$$

donde $g_m[n] = \|x_m[n]\|^2$ y $\overline{g_m[n]}$ denota valor medio. Como se discute en [11] esta elección confiere mayor robustez al método de regresión.

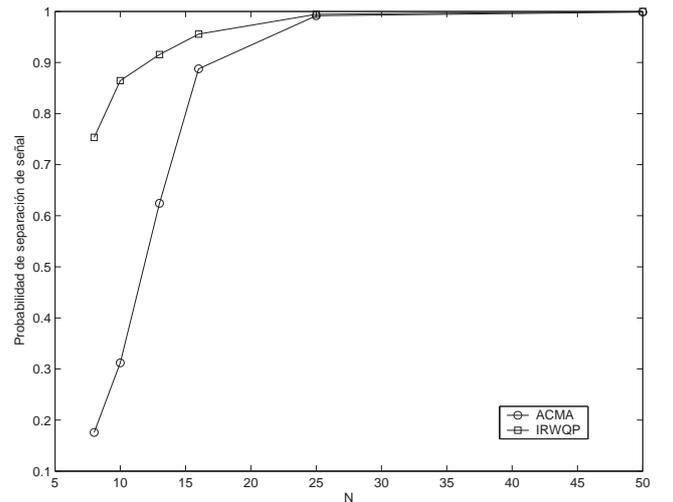


Fig. 1. Probabilidad de extracción correcta en un ejemplo con tres señales QPSK y una de módulo no constante, SNR = 15dB; en función del número de observaciones N .

En la Figura 1 se compara la probabilidad de extracción correcta de señal para el método basado en SVM y el ACMA, en función del número de observaciones. El criterio utilizado para calificar una señal extraída como correcta consiste en el cálculo de la relación entre el error cuadrático de la señal respecto a la original y la potencia de la señal original, considerando acertada la separación si se cumple

$$\frac{\|s_k - \tilde{y}_k\|_2^2}{\|s_k\|_2^2} > 0.5, \quad (26)$$

donde s_k es la señal QPSK original, e \tilde{y}_k es una versión del vector y_k escalado por el complejo que minimiza el error cuadrático con respecto a s_k . La aplicación de este factor es necesaria debido a que existe una ambigüedad en la fase de la señal recuperada.

En la Figura 1 se observa cómo el método propuesto ofrece una mayor probabilidad de extracción principalmente cuando el número de observaciones es muy reducido ($N < 25$).

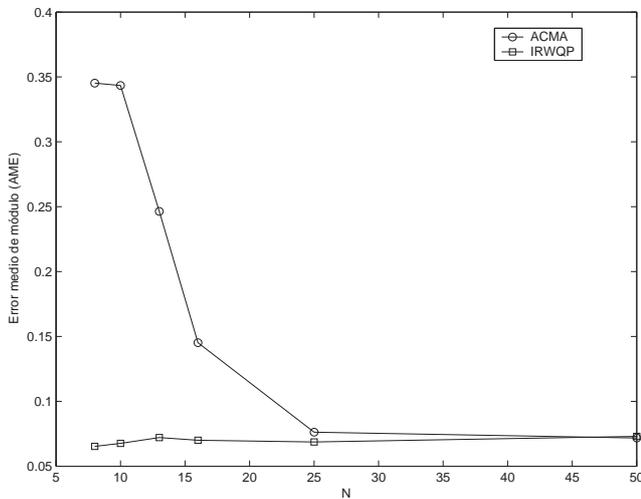


Fig. 2. Error medio en el módulo (AME) en un ejemplo con tres señales QPSK y una de módulo no constante, SNR = 15dB; en función del número de observaciones N .

Para las señales extraídas correctamente se ha obtenido, además, como figura de mérito el error medio en el módulo “Average Modulus Error” (AME), definido como

$$\left[\frac{1}{N} \sum_{n=0}^{N-1} (||y_k[n]|| - 1)^2 \right]^{1/2}. \quad (27)$$

De nuevo se observa en la Figura 2 cómo el método ACMA pierde efectividad para señales con $N < 25$, mientras que algoritmo basado en SVM consigue valores del AME menores que 0.1 en todos los casos simulados. Hay que hacer notar, sin embargo, que la mejora en los resultados se consigue a cambio de un incremento en el coste computacional, aspecto en el que el ACMA es más ventajoso.

VII. CONCLUSIONES

En este artículo el problema de conformación ciega de haz para señales de módulo constante se ha formulado como un

problema de regresión y se ha aplicado la técnica basada en SVM para resolverlo. Se ha propuesto un algoritmo iterativo que converge a una de las señales de módulo constante presente en las observaciones. Para la extracción de varias señales de módulo constante se ha elaborado un método con varios conformadores actuando en paralelo en el que la señal extraída por cada conformador es cancelada de la señal a procesar por el siguiente, convergiendo así cada uno de ellos a una de las distintas señales presentes.

Las simulaciones realizadas muestran que el método propuesto ofrece un mejor comportamiento que el método ACMA, principalmente cuando el número de observaciones es muy reducido. El método propuesto, no obstante, requiere una elevada carga computacional. La aplicación de técnicas eficientes recientemente propuestas para resolver problemas de programación cuadrática (QP) pueden ser de utilidad para reducir este inconveniente [12], [13].

AGRADECIMIENTOS

Este trabajo ha sido parcialmente financiado a través del proyecto TIC2001-0751-C04-03, del Ministerio de Ciencia y Tecnología.

REFERENCES

- [1] B. D. Van Veen, K. M. Buckley, *Beamforming: A Versatile Approach to Spatial Filtering*, IEEE Signal Proc. Magazine, vol.7, no. 3, pp. 4-24, April 1988 ASSP-31, pp. 459-471, Apr. 1983.
- [2] J. R. Treichler, B. G. Agee, *A New Approach to Multipath Correction of Constant Modulus Signals*, IEEE Trans. on Acoust., Speech, Signal Processing, vol. ASSP-31, pp. 459-471, Apr. 1983.
- [3] R. Gooch, J. Lundell, *The CM Array: An Adaptive Beamformer for Constant Modulus Signals*, Proc. ICASSP, pp. 2523-2526, Tokyo, 1986.
- [4] J. Shynk, R. Gooch, *The Constant Modulus Array for Cochannel Signal Copy and Direction Finding*, IEEE Trans. on Signal Processing, vol. 44, no. 3, pp. 777-787, March 1996.
- [5] B. G. Agee, *The Least-Squares CMA: A New Technique for Rapid Correction of Constant Modulus Signals*, in Proc. ICASSP, pp. 953-956, Tokyo, 1986.
- [6] A. J. van der Veen, A. Paulraj, *An Analytical Constant Modulus Algorithm*, IEEE Trans. on Signal Processing, vol. 44, no. 5, May 1996.
- [7] I. Santamaría, C. Pantaleón, L. Vielva, J. Ibáñez, *Blind Equalization of Constant Modulus Signals via Support Vector Regression*, in Proc. ICASSP, Hong Kong, China, 2003.
- [8] I. Santamaría, C. Pantaleón, L. Vielva, J. Ibáñez, *Blind Equalization of Constant Modulus Signals Using Support Vector Machines*, IEEE Trans. on Signal Processing, (en revisión), 2003.
- [9] P. W. Holland, R. E. Welch, *Robust Regression Using Iterative Reweighted Least Squares*, Commun. Statist. Theory Methods, vol. A 6, n° 9, pp. 813-827, 1997.
- [10] S. R. Gunn, *Matlab Support Vector Machine Toolbox*, disponible en <http://www.isis.ecs.soton.ac.uk/isystems/kernel/>, University of Southampton, Image Speech and Intelligent Systems Research Group, UK, 1998.
- [11] V. Cherkassky, Y. Ma, *Selection of Meta-parameters for Support Vector Regression*, in J. R. Dorronsoro (Ed.), Proc. of ICANN'2002, pp. 687-693, Springer-Verlag, Berlin, 2002.
- [12] J. Platt, *Fast Training of Support Vector Machines Using Sequential Minimal Optimization*, in B. Schölkopf, C. J. C. Burges and A. J. Smola, (Eds.), *Advances in Kernel Methods: Support Vector Learning*, pp. 185-208, The MIT Press, Cambridge, MA, 1999.
- [13] T. Joachims, *Making Large-Scale Support Vector Machine Learning Practical*, en B. Schölkopf, C. J. C. Burges and A. J. Smola, (Eds.), *Advances in Kernel Methods: Support Vector Learning*, pp. 169-184, The MIT Press, Cambridge, MA, 1999.