Estimación recursiva simultánea de potencia y DOA por ajuste de la matriz de covarianza asociada a la apertura.

A. Perez-Neira, M.A. Lagunas
TSC Dept., E.T.S.I. Telecomunicación Barcelona
c/ Sor Eulalia de Anzizu s/n. Apdo. 30002
08080 Barcelona (España).

Abstract: The small error approximation is used to derive a linear relationship between the source parameters (i.e. power levels and directions of arrival or DOA's) and a measurement of the covariance error matrix, defined as the difference between a nonparametric consisten estimate of the spectral density matrix and a covariance model from the scenario parameters (power level and directions of arrival). The resulting framework allows the disign of a recursive algorithm which provides a simultaneous and adaptive estimation of the sources parameters no matter what is the source waveform or modulation. Due to its similarities with the Kalman filter structure [1], good performance is expected, mainly in the presence of sensors' misfunctioning, low signal-to-noise ratio (SNR), etc...

1.INTRODUCCION.

El estudio de técnicas para la obtención de los diversos parámetros que caracterizan las ondas incidentes en una apertura de sensores es un tema constantemente presente en la literatura que trata sobre procesado de "arrays". Una técnica que conduce a resultados óptimos es la de Máxima Verosimilitud (ML), sin embargo, debido a la alta carga computacional que conlleva, en la práctica se emplean métodos, que aunque subóptimos, son más ágiles de cálculo. Entre estos los más ampliamente empleados son los Métodos de Alta Resulución (p.ej.:Music, Estimación de Capón), capaces de tratar situaciones en las que las fuentes de información están muy cercanas. El inconveniente de estos métodos es que son poco robustos y fallan si la SNR de las señales en los sensores es baja o si se dispone de poco número de muestras para procesar las señales recibidas. Además, si bien la información que proporcionan sobre la localización de las fuentes es satisfactoria, no sucede lo mismo cuando se trata de la información espectral de potencia. Con el propósito de obtener resultados más completos, se suele recurrir a técnicas que estiman los parámetros en varios pasos: a través de métodos de alta resolución se obtiene una primera aproximación que luego es mejorada mediante procesos iterativos (Ligget [2],Böhme [3]). Otros procedimientos son los que han trabajan directamente sobre el criterio de máxima verosimilitud, agilizándolo mediante el empleo de técnicas de gradiente, pero, debido a que lo que realizan estas técnicas es una búsqueda de mínimos locales, están limitadas por la necesidad de disponer de una correcta inicialización de los parámetros a estimar [4,5].

En este trabajo, se propone una técnica que pretende evitar los mencionados problemas de elevado coste computacional y de necesidad de una adecuada inicialización de los parámetros a estimar.

2.Estimación de máxima verosimilitud y ajuste de la matriz de covarianza asociada a la apertura.

La señal a la salida de una determinada agrupación de sensores puede modelarse en el dominio frecuencial del siguiente modo:

$$\underline{X}(w) = \underline{\underline{A}}(w) \,.\, \underline{S}(w) + \underline{U}(w) \qquad (1)$$

en donde $\underline{\underline{A}}(w)$ es la matriz fomada por los vectores de dirección de llegada $\underline{a}_{sn}$, que contienen la información de los ángulos de llegada de las señales incidentes, $\underline{S}(w)$ es la transformada de Fourier de las señales que llegan y $\underline{U}(w)$ es el ruido blanco de naturaleza gaussiana que se haya en los sensores y que suponemos incorrelado con las señales de información. La matriz espectral de potencia de la salidá de la apertura se puede expresar como:

$$\underline{\underline{C}}x(w) = \underline{\underline{A}}(w) \,.\, \underline{\underline{C}}s(w) \,.\, \underline{\underline{A}}^H(w) + \sigma^2 . \underline{\underline{I}} \qquad (2)$$

en donde $\underline{\underline{C}}s(w)$ es la matriz espectral de las señales y $\sigma^{\wedge 2}$.es la potencia de ruido.

Recurriendo a las propiedades asintóticas que posee la transformada de Fourier de la salida del array en (1), la función de máxima verosimilitud puede formularse como se indica a continuación:

$$L(\underline{E}) = -\log \det \underline{\underline{C}}x(\underline{E}) - tr\,(\hat{\underline{\underline{C}}}_x \cdot \underline{\underline{C}}x(\underline{E})^{-1}) \qquad (3)$$

en donde $\underline{\underline{C}}x(\underline{E})$ es la matriz de covarianza estimada y $\hat{\underline{\underline{C}}}_x$ es una estimación estadísticamente consistente de la matriz de covarianza real:

$$\hat{\underline{\underline{C}}}_x .= (1/K).\Sigma_{k=1}^{N} \underline{X}k \,.\, \underline{X}k^H \qquad (4)$$

el vector E contiene los parámetros a estimar: la matriz espectral de las señales,$\underline{\underline{C}}s$, los DOA´s,ξ, y la potencia del ruido, σ^2: $\underline{E}^T=(\underline{\underline{C}}s,\xi\ ,\sigma^2)$.

En la notación de (3), se ha omitido el parámetro "w" ya que nuestro trabajo se centrará en el caso de banda estrecha.

Para construir el estimador de máxima verosimilitud se maximiza primeramente la función $L(\underline{\underline{C}}s,\xi\ ,\sigma^2)$ sobre $\underline{\underline{C}}s$ and σ^2, manteniendo ξ fijo. En el caso de obtener soluciones explícitas éstas se sustituyen dentro de la función de verosimilitud (3) y se maximiza el criterio resultante sobre ξ. La estimación ML resultante es la presentada en [3]; que para la matriz espectral de potencia , $\underline{\underline{C}}s$, se puede formular:

$$.\hat{C}_s\,(x) = \underline{\underline{P}} \cdot \hat{\underline{\underline{C}}}_x \cdot \underline{\underline{P}} - \hat{\ }\sigma^2.(\underline{\underline{A}}^H . \underline{\underline{A}})^{-1} \qquad (5.a)$$

en donde la $\underline{\underline{P}}$ es la matriz de proyección en el subespacio de direcciones de llegada:

$$\underline{\underline{P}} = \underline{\underline{A}} \cdot (\underline{\underline{A}}^H . \underline{\underline{A}})^{-1} . \underline{\underline{A}}^H \qquad (5.b);$$

Para la potencia de ruido la estimación final es la que presentamos en (6):

$$\hat{\ }\sigma^2 = tr\ \{\ (\underline{\underline{I}}-\underline{\underline{P}}) \cdot \hat{\underline{\underline{C}}}_x\ \} \qquad (6)$$

Tal y como se muestra en [6] estos resultados de la estimación ML son los mismos que los obtenidos si el criterio que se emplea para realizar la estimación es el de minimizar el error cuadrático de la matriz de covarianza, es decir, ajustar la matriz de covarianza asociada a la apertura, minimizando el error o diferencia entre la covarianza de la señal observada y la de la señal estimada:

$$Q(\underline{E}) = tr\ \{(\underline{\underline{C}}x(\underline{E}) - \hat{\underline{\underline{C}}}_x)^2\ \} \qquad (7)$$

Este último criterio es utilizado en muchas ocasiones por el buen compromiso que presenta entre la carga de cálculo que conlleva y carácter óptimo de los resultados. Como ejemplo mencionamos el trabajo realizado por d'Assumpcao en [7], el autor diseña diferentes estimadores de la potencia de las señales emitidas por las fuentes empleando el criterio de mínimo error cuadrático en (7). Del mismo modo sucede que, en la mayoría de las situaciones, la medida del error que se comete al estimar la matriz de covarianza es necesaria cuando se pretende conocer las características de los diferentes frentes de ondas presentes en un escenario y la herramienta utilizada es el procesado de los datos recogidos por una apertura. El modo en el que se mida el error de covarianza dependerá del problema a tratar [8]. Por ejemplo, en el caso de que en lugar de una matriz de error se necesite un escalar, la elección más correcta será considerar la norma de la matriz de error, generalmente la traza. En nuestro trabajo, se verá en la sección 4 como, para el diseño del filtro lineal que ha de estimar los parámetros de la fuentes, la medida del error que tomaremos deberá cumplir dos requisitos: por una parte deberá tratarse de un vector de error, por la otra, deberá depender linealmente de los parámetros de las fuentes recogidos en el vector $\underline{E}$. En la siguiente sección se expone el modo de obtener una medida de error que se ciña a dichas características.

3.Medida del error de covarianza: aproximación lineal

Consideremos un escenario formado por N fuentes incoherentes, la estimación del vector de parámetros en la interación n $\hat{\underline{E}}_n$, estará formada por cada uno de los parámetros estimados, potencia de las fuentes y del ruido y localización de las fuentes:

$$\hat{\underline{E}}_n = (\hat{c}_{1n}, \hat{c}_{2n}, \ldots, \hat{c}_{Nn}, \hat{\xi}_{1n}, \hat{\xi}_{2n}, \ldots, \hat{\xi}_{Nn}, \hat{\sigma}^2)^T \qquad (8)$$

La matriz de error expresada en función de estas magnitudes será:

$$\tilde{\underline{\underline{C}}}_x = \underline{\underline{C}}_x - \underline{\underline{C}}_x(\hat{\underline{E}}_n) = \Sigma_{s=1}^{N} (\mathbf{c_s \cdot \underline{a}_{sn} \cdot \underline{a}_{sn}^H - \hat{c}_{sn} \cdot \underline{a}(\hat{\xi}_{sn}) \cdot \underline{a}^H(\hat{\xi}_{sn})}) + \tilde{\sigma}^2 \cdot \underline{\underline{I}} \qquad (9)$$

Si nos ceñimos a la parte marcada en negrita de la expresión anterior, ésta puede reescribirse como un sumatorio de los términos siguientes:

$$\tilde{c}_{sn}\, \underline{a}_{sn} \cdot \underline{a}_{sn}^H - \hat{c}_{sn} \cdot (\underline{a}_{sn} \cdot \underline{a}_{sn}^H - \hat{\underline{a}}_{sn} \cdot \hat{\underline{a}}_{sn}^H) \qquad (10)$$

en donde el error del nivel de potencia de las fuentes se ha expresado por $\tilde{c}_{sn}$:

$$\tilde{c}_{sn} = c_{sn} - \hat{c}_{sn} \qquad (11)$$

Es en este punto cuando realizaremos una aproximación lineal de error basándonos en la suposición de que el error que se comete en la estimación es muy pequeño. De este modo, la medida del error cometido en la estimación de la matriz de covarianza se podrá expresar linealmente dependiente del error de parámetros. Por lo tanto, si se asume que, en una iteración dada $\hat{\underline{a}}_{sn}$ se haya muy próximo al vector de direcciones real, el término escrito en (10) puede aproximarse por (12):

$$\tilde{c}_{sn}\, \hat{\underline{a}}_{sn} \cdot \underline{a}_{sn}^H + \hat{c}_{sn} \cdot (\tilde{\underline{a}}_{sn} \cdot \underline{a}_{sn}^H) \qquad (12)$$

en donde $\tilde{\underline{a}}_{sn}$ representa el error cometido en la estimación del vector de dirección de llegada correspondiente a la fuente s:

$$\tilde{\underline{a}}_{sn} = \underline{a}_{sn} - \underline{a}(\hat{\xi}_{sn}) = \underline{a}_{sn} - \hat{\underline{a}}_{sn} \qquad (13)$$

Con el propósito de conseguir una relación lineal entre el error medido, (13), y el error cometido al estimar el ángulo de elevación se vuelve a emplear la aproximación de pequeño error en cada una de las componentes "q" del vector de error de dirección de llegada, $\tilde{\underline{a}}_{sn}$

$$(\tilde{\underline{a}}_s)_q = (\underline{a}_s)_q - (\hat{\underline{a}}_s)_q = \exp(k_q \cdot \mathrm{Sin}\,(\xi_s)) - \exp(k_q \cdot \mathrm{Sin}\,(\hat{\xi}_s)) =$$

$$= \exp(k_q \cdot \mathrm{Sin}\,(\tilde{\xi}_s + \hat{\xi}_s)) - \exp(k_q \cdot \mathrm{Sin}\,(\hat{\xi}_s)) \sim \exp(k_q \cdot (\mathrm{Sin}\,\hat{\xi}_s + \tilde{\xi}_s \cdot \mathrm{Cos}\hat{\xi}_s)) -$$

$$- \exp(k_q \cdot \mathrm{Sin}\,(\hat{\xi}_s)) \sim \exp(k_q \cdot \mathrm{Sin}\,(\hat{\xi}_s)) \cdot (j \cdot k_q \cdot \mathrm{Cos}(\hat{\xi}_s) \cdot \tilde{\xi}_s)) \qquad (14)$$

$$(\tilde{\underline{a}}_s)_q \sim (\hat{\underline{a}}_s)_q \cdot (j \cdot k_q \cdot \mathrm{Cos}\hat{\xi}_s \cdot \tilde{\xi}_s) \qquad (15)$$

en donde $k_q=(2.\pi.f/c).dq.\mathrm{Cos}(\phi_q)$, siendo "f" la frecuencia central y d_q y ϕ_q son la distancia al centro de fase y el ángulo de azimuth del sensor "q" respectivamente (se supone una apertura lineal con los sensores uniformemente distribuidos).

Concluyendo, el error cometido al estimar la matriz de covarianza puede expresarse como una combinación lineal del error de parámetros:

$$\underline{\underline{\tilde{C}}}_x = .\Sigma_{s=1}^{N} (\tilde{c}_{ns} \cdot \hat{\underline{a}}_{ns} \cdot \hat{\underline{a}}_{sn}^{H} + \tilde{\xi}_{sn} \cdot \underline{a}'_{sn} \cdot \hat{\underline{a}}_{sn}^{H}) + \tilde{\sigma}^2 \cdot \underline{\underline{I}} \quad (16)$$

en donde cada componente del vector $(\underline{a}'_{sn})_q$ tiene la forma:

$$(\underline{a}'_{sn})_q = (\hat{\underline{a}}_{sn})_q \cdot (k_q \cdot \mathrm{Cos}\hat{\xi}_s) \cdot \hat{c}_{ns} \quad (17)$$

La medida del error de covarianza que se empleará para diseñar el filtro lineal ha de estar expresado en forma vectorial, es por ello que se tomará la primera columna de la matriz de error que ha resultado en (16). Ello equivale a tomar un sensor como referencia y minimizar la diferencia entre la correlación cruzada real y la medida que se obtiene al correlar la señal en el sensor de referencia y el resto de sensores que forman la apertura. La formaulación de esta medida será:

$$\underline{\varepsilon}_n = \underline{\underline{\tilde{C}}}_x \cdot \underline{1} = .\Sigma_{s=1}^{N} (\tilde{c}_{sn} \cdot \hat{\underline{a}}_{sn} + \tilde{\xi}_{sn} \cdot \underline{a}'_{sn}) + \tilde{\sigma}^2 \cdot \underline{1} \quad (18.a); \qquad \underline{1} = (1,0,0,...0)^T \quad (18.b)$$

Este último resultado es básico para poder obtener el correspondiente filtro lineal, ya que establece la relación fundamental entre el error medido y el error de parámetros:

$$\underline{\varepsilon}_n = \underline{\underline{H}}_n \cdot \underline{\tilde{E}}_n \quad (19.a)$$

siendo: $$\underline{\tilde{E}}_n = (\tilde{c}_{1n}, \tilde{c}_{2n}, ...\tilde{c}_{Nn}, \tilde{\xi}_{1n}, \tilde{\xi}_{2n}, ..., \tilde{\xi}_{Nn}, \tilde{\sigma}^2)^T \quad (19.b)$$

$$\underline{\underline{H}}_n = (\hat{\underline{a}}_{1n}, \hat{\underline{a}}_{2n}, ..., \hat{\underline{a}}_{Nn}, \underline{a}'_{1n}, \underline{a}'_{2n}, ..., \underline{a}'_{Nn}, \underline{1})^T \quad (19.c)$$

4.Filtro adaptativo lineal.

Por la brevedad que es exigida en esta presentación, sólo expondremos un resumen del procedimiento empleado para llevar a cabo nuestro objetivo de estimar simultánea y adaptativamente la información de potencia y de localización de las fuentes. Las simulaciones que se presentarán forman parte de un conjunto completo de resultados de los cuales se concluye el comportamiento adaptativo del algoritmo y lo satisfactorio de sus resultados.

La actualización de la estimación de los parámetros se realiza a través de la matriz $\underline{\underline{K}}$:

$$\underline{\hat{E}}_{n+1} = \underline{\hat{E}}_n + \underline{\underline{K}}_n \cdot \underline{\varepsilon}_n \quad (20)$$

Si se asume que el vector que contiene los parámetros reales $\underline{E}$ permanece estacionario con las iteraciones más un cierto ruido de innovación de matriz de covarianza $\underline{Q}$, que se considera diagonal, la ortogonalidad entre el error de parámetros $\underline{\tilde{E}}_{n+1}$ y el error de medida $\underline{\varepsilon}_n$ proporciona la matriz de ganancia óptima:

$$\underline{\underline{K}}_n = \underline{\underline{\Sigma}}_n \cdot \underline{\underline{H}}_n^H \cdot (\underline{\underline{H}}_n \cdot \underline{\underline{\Sigma}}_n \cdot \underline{\underline{H}}_n^H)^{-1} = \underline{\underline{\Sigma}}_n \cdot \underline{\underline{H}}_n^H \cdot \underline{\underline{\zeta}}_n^{-1} \quad (21)$$

La matriz $\underline{\underline{\Sigma}}_n$ se define como la covarianza del error de parámetros en la iteración n. La actualización de dicha matriz se basa por tanto en la evolución estacionaria del vector que contiene los parámetros a estimar y en la covarianza del error de innovación. La actualización que resulta es la siguiente:

$$\underline{\underline{\Sigma}}n+1 = \underline{\underline{\Sigma}}n - \underline{\underline{K}}n \,.\, \underline{\underline{\zeta}}n \,.\, \underline{\underline{K}}n^{H} + \underline{\underline{Q}} \qquad (22)$$

Como ejemplo del comportamiente de máxima verosimilitud que presenta este algoritmo de rápida convergencia se muestran en la figura 1 los resultados en el caso de que el número de fuentes presentes en el escenario sea 2 y ello no coincida con el número de fuentes que se ha indicado al algoritmo que ha de estimar (1 fuente). La estimación resultante es aproximadamente la media aritméťica de los parámetros reales. En cuanto al carácter adaptativo del filtro, en la simulación de la figura 2 se aprecia cómo el algoritmo diseñado ofrece un buen seguimiento de la fuente móvil (de 10º a -10º) a pesar de la baja relación señal a ruido con la que llega la señal emitida a los sensores (-3 dB´s) y a pesar de la pequeña apertura tomada. En ambas figuras, la apertura considerada estaba formada por 7 sensores y la matriz de covarianza $\hat{\underline{\underline{C}}}_X$, que en el algoritmo se emplea como matriz de covarianza real, se calculó a partir de únicamente 500 muestras temporales o "snapshots".

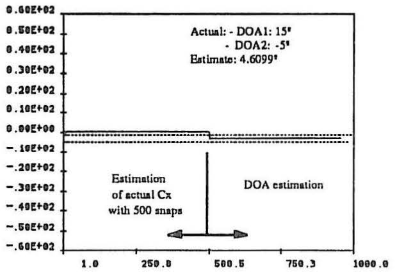

Fig. 1

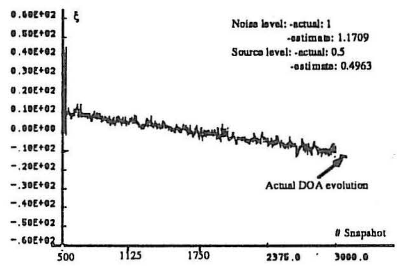

Fig.2.

5.Conclusiones finales.

A través del ajuste de la matriz de covarianza asociada a la apetura mediante la aproximación de pequeño error se ha presentado cómo el error cometido en la estimación de la matriz de covarianza puede formularse linealmente dependiente del vector de error de parámetros. Ello ha sido la base del diseño de un filtro adaptativo lineal que permita estimar simultáneamente de forma rápida y con un comportamiento cercano al de máxima verosimilitud los parámetros de las fuentes , incluso en el caso de que la matriz de covarianza real sea calculada únicamente apartir de una pequeña cantidad de muestras temporales. Las numerosas simulaciones realizadas muestran como la baja carga computacional del algoritmo así como su capacidad de seguimiento de fuentes lo configuran como un buen método para aplicaciones de conformación de haz con referencia espacial. Por otra parte, el estimador diseñado presenta resultados buenos y robustos aún en el caso de fuentes próximas, independientemente de cúal sea el vector inicial de parámetros del que se parta para hacer la estimación. En definitiva, posee toda una serie de características que no están presentes en muchos de los métodos de estimación hasta ahora existentes.

6.Referencias

[1] B.D.O. Anderson and J.B. Moore, "Optimal Filtering", Englewood Cliffs, N.J. Prentice-Hall, 1979.

[2] W.S.Liggett,"Passive Sonar:Fitting Models to Multiple Time Series", in Signal Processing, Griffiths, J.W.R. et al. eds., Academic Press, New York, 1973, pp.327-345.

[3] J.F.Böhme,"Estimation of Source Parameters by Maximum Likelihood and Nonlinear Regression",IEEE proceedings 1984, 7.3.1-7.3.4.

[4] Y.Bresler, A.Macovski, "Exact Maximum Likelihood Parameter Estimation of Superimposed Exponential Signals in Noise", IEEE Trans.1986, Vol. ASSP-34, no.5, pp.1087-1089.

[5] H.Watanabe, M.Suzuki,"Maximum Likelihood Bearing Estimation by Quasi-Newton Method using a Uniform Linear Array", IEEE proceeding 1991, p.p. 3325-3328.

[6] D.Kraus, G.Schmitz, J.F.Böhme, "Least Squares Estimates for Source Location and Asymptotic Behaviours", in Signal Processing V , Elsevier Science Publishers, 1990, p.p. 649-652

[7] H.A. d'Assumpcao, "Some new Signal Processors for Arrays of Sensors", IEEE Transactions on Information Theory, Vol. IT-26, no.4, July 1980, p.p.441-454.

[8] M.Basseville, "Distance Measures for Singal Processing and Pattern Recognition", Signal Processing 1989, Elsevier Science Publishers, p.p. 349-369.