

$$\begin{aligned}
 \mathbf{D} &= [d_{i\nu}], \quad d_{i\nu} = M_\nu(x_i) \quad i = 1, \dots, N_D \quad \nu = 1, \dots, M \\
 \bar{b} &= [b_i], \quad b_i = M(x_i) \quad i = 1, \dots, N_D \\
 \bar{c} &= [C_\nu], \quad \nu = 1, \dots, M
 \end{aligned} \tag{7}$$

quedando la ecuación (5) escrita

$$\mathbf{D}\bar{c} = \bar{b}. \tag{8}$$

Los coeficientes $C_\nu \geq 0$ representan las proporciones en que las distintas tierras raras se encuentran en la mezcla. Para determinarlos en base a un conjunto de N_D medidas experimentales, obtenidas con cierto error, presentaremos un algoritmo basado en el Principio de Máxima Entropía. Introduciremos en la siguiente sección algunos elementos que utilizaremos para ello.

ALGUNOS CONCEPTOS BÁSICOS SOBRE LA TEORÍA DE LA INFORMACIÓN Y EL PRINCIPIO DE MÁXIMA ENTROPÍA

En 1948 Claude Shannon⁵ desarrolló un esquema matemático que permite cuantificar la información contenida en una distribución de probabilidades $\{p_j; j = 1, \dots, M\}$ definida sobre un conjunto de M sucesos complementarios. Dentro de este esquema, la medida de información está dada por la expresión:

$$S = -k \sum_{j=1}^M p_j \ln p_j \tag{9}$$

donde k es la unidad de información. Esta función S , conocida también con el nombre de "Entropía Estadística", representa la cantidad de información que, en valor medio, recibe un observador al enterarse de la ocurrencia de algún suceso entre los M posibles. Esta función es, por lo tanto, una medida de lo que el observador "ignora" antes de conocer el resultado de un experimento. En 1957 Jaynes⁶ adoptó esta definición de Shannon para operar matemáticamente con el concepto de "ignorancia" y propuso un método de inferencia, conocido como Principio de Máxima Entropía, por medio del cual dio una prescripción para seleccionar una distribución de probabilidades en base al menor número de hipótesis. Sintéticamente, el PME establece que:

Si se dispone de un número incompleto de datos $\{O_i; i = 1, \dots, N < M\}$ dados en valor medio, es decir:

$$\langle O_i \rangle = \sum_{j=1}^M p_j O_{ji} \quad ; \quad i = 1, \dots, N, \tag{10}$$

de todas las distribuciones de probabilidades que reproducen los datos conocidos, la que posee mayor poder de inferencia es la que además maximiza la ignorancia del observador.

$$\mathbf{D}^T \bar{\mathbf{b}}^0 = [F_k] ; F_k = \sum_{i=1}^{N_D} d_{ki} b_i^0, \quad (18)$$

Resolver el sistema (16) es equivalente a dar una solución de cuadrados mínimos, pero para el problema que estamos tratando esta solución no es la adecuada fundamentalmente por dos motivos:

- 1- La restricción de cuadrados mínimos por sí misma no garantiza que la solución pertenezca al dominio admitido $C_\nu > 0$; ($\nu = 1, \dots, M$).
- 2- No tiene en cuenta explícitamente el error de los datos.

Para salvar estos inconvenientes proponemos aquí un algoritmo que resuelve iterativamente el sistema (16) dentro del error con que se conocen las ecuaciones. En primer término reescribimos los elementos de $\bar{\mathbf{c}}$ en la forma:

$$C_\nu = A p_\nu ; p_\nu \geq 0 ; \nu = 1, \dots, M, \quad (19)$$

donde $A > 0$ es una constante tal que:

$$\sum_{\nu=1}^M p_\nu = 1, \quad (20)$$

de modo que el sistema de ecuaciones a resolver resulta:

$$A \mathbf{D}^T \mathbf{D} \bar{\mathbf{p}} = \mathbf{D}^T \bar{\mathbf{b}}^0. \quad (21)$$

Con la condición de normalización impuesta, los pesos p_ν definen un espacio de probabilidades sobre un conjunto discreto de M sucesos, cada uno de los cuales corresponde a una tierra rara particular. Esta distribución de probabilidades será la que caracterice al sistema en estudio, ya que el mismo es de clara naturaleza estadística: aquí la muestra (sistema) está compuesta por distintos subsistemas (diferentes tierras raras), cada uno con características definidas y poblando la muestra en distintas proporciones. Si existiera la posibilidad de realizar algún "hipotético" experimento sobre este espacio de sucesos, como por ejemplo aislar algún subsistema elegido al azar y medir su magnetización, la información que se ganaría al conocer el resultado de tal hipotético experimento puede ser calculada, en valor medio, utilizando la entropía propuesta por Shannon⁵

$$S = -k \sum_{\nu=1}^M p_\nu \ln p_\nu. \quad (22)$$

Sin embargo, la ecuación (5) da cuenta de que ese hipotético experimento en nuestro caso no puede realizarse y que, como ha sido insistentemente enfatizado por E. Jaynes³, nuestras medidas pertenecen a otro espacio y nos proporcionan únicamente **en valor medio** el resultado de las correspondientes medidas en el espacio de sucesos. Sólo existe una situación en la que la ecuación (5) sí daría una medida en el espacio de sucesos y es precisamente el caso trivial, en el que $p_\nu = 1$ para algún ν y 0 para el resto. En este

Supongamos que el máximo ϵ_k corresponde a $k = k_1$. A los pesos $p_\nu^{(1)}$ los elegimos ahora de manera tal que maximicen la entropía S sujeta a la restricción:

$$\sum_{\nu=1}^M p_\nu^{(1)} O_{\nu k_1} = 0, \quad (29)$$

lo que es equivalente a exigir que se satisfaga la k_1 -ésima ecuación del sistema (21). De acuerdo con lo expuesto en la sección anterior, los $p_\nu^{(1)}$ que resultan son de la forma:

$$p_\nu^{(1)} = \frac{\exp(-\lambda_1 O_{\nu k_1})}{\sum_{s=1}^M \exp(-\lambda_1 O_{s k_1})} \quad (30)$$

donde el multiplicador λ_1 surge de resolver la ecuación (29).

Con los $p_\nu^{(1)}$ podemos construir las predicciones $F_k^{(1)}$ ($k = 1, \dots, M$) y el correspondiente nuevo conjunto de ϵ_k . Después de seleccionar el máximo elemento de $\{\epsilon_k\}$, sea éste ϵ_{k_2} , obtenemos los $p_\nu^{(2)}$ maximizando la entropía sujeta a dos restricciones, impuestas por el cumplimiento de las ecuaciones correspondientes a $k = k_1$ y $k = k_2$. En general para la aproximación de orden J se obtiene:

$$p_\nu^{(J)} = \frac{\exp(-\sum_{t=1}^J \lambda_t O_{\nu k_t})}{\sum_{s=1}^M \exp(-\sum_{t=1}^J \lambda_t O_{s k_t})}, \quad (31)$$

donde los multiplicadores λ_t ; $t = 1, \dots, J$ surgen de resolver las J ecuaciones:

$$\sum_{\nu=1}^M p_\nu^{(J)} O_{\nu k_i} = 0 \quad ; \quad i = 1, \dots, J. \quad (32)$$

Este proceso iterativo puede ser finalizado cuando se verifica que:

$$\epsilon_k \leq \frac{\Delta F_k}{|F_k|} \quad ; \quad k = 1, \dots, M, \quad (33)$$

donde los ΔF_k quedan determinados por los errores de los datos Δb_i^0 ; $i = 1, \dots, N_D$,

$$\Delta F_k = \sum_{i=1}^{N_D} d_{ki} \Delta b_i^0 \quad ; \quad k = 1, \dots, M, \quad (34)$$

Supongamos que la convergencia (32) se logra en la L -ésima iteración. Es importante resaltar que la solución $C_\nu^{(L)}$; $\nu = 1, \dots, M$ no minimiza el funcional Φ definido por (13), sino que garantiza que el mismo esté acotado por la desigualdad:

$$\sum_{i=1}^{N_D} (b_i^0)^2 - \sum_{k=1}^M C_k^{(L)} (F_k + \Delta F_k) \leq \Phi \leq \sum_{i=1}^{N_D} (b_i^0)^2 - \sum_{k=1}^M C_k^{(L)} (F_k - \Delta F_k) \quad (35)$$

Con los coeficientes $C_\nu^{(L)}$ se puede inferir el valor de cada una de las medidas

ν	Ión	S_ν	L_ν	J_ν	C_ν	C_ν^*	$C_\nu^{(1)}$	C_ν^{**}
1	Ce^{3+}	$\frac{1}{2}$	3	$\frac{5}{2}$	0.003	7.001	0.018	0.000
2	Pr^{3+}	1	5	4	0.007	-10.42	0.013	0.005
3	Nd^{3+}	$\frac{1}{2}$	6	$\frac{4}{2}$	0.006	-10.49	0.012	0.000
4	Pm^{3+}	2	6	4	0.014	-2.659	0.015	0.000
5	Gd^{3+}	$\frac{7}{2}$	0	$\frac{7}{2}$	0.105	-13.39	0.084	0.000
6	Tb^{3+}	3	3	6	0.199	8.869	0.165	0.000
7	Dy^{3+}	$\frac{5}{2}$	5	$\frac{15}{2}$	0.258	-4.881	0.264	0.226
8	Ho^{3+}	2	6	8	0.193	-3.341	0.245	0.000
9	Er^{3+}	$\frac{3}{2}$	6	$\frac{15}{2}$	0.149	10.78	0.126	0.737
10	Tm^{3+}	1	5	6	0.057	-2.702	0.041	0.000
11	Yb^{3+}	$\frac{1}{2}$	3	$\frac{7}{2}$	0.007	21.29	0.015	0.000

Tabla I. C_ν : Pesos conocidos. C_ν^* : Pesos obtenidos por cuadrados mínimos sin restricción de positividad. $C_\nu^{(1)}$: Pesos obtenidos con el algoritmo propuesto. C_ν^{**} : Pesos obtenidos por cuadrados mínimos con restricción de positividad.

problema en cuestión. Como se ve en la tabla, para datos con 1% de error varios de los coeficientes que resultan son negativos, por lo que no pueden ser tomados como válidos por no satisfacer la restricción de positividad de la solución buscada.

Utilizando un algoritmo de (CM) que sí incorpora la restricción de positividad¹⁰, los coeficientes que resultan son los que se muestran en la última columna de la tabla. Como se ve, la mayoría de estos coeficientes tienen valor cero, resultado que no es sorprendente pues la teoría de optimización convexa asegura que si el mínimo absoluto de un funcional convexo cae fuera de la región admisible, que en nuestro caso es la región determinada por la intersección de los hiperplanos

$$C_\nu \geq 0 ; \quad \nu = 1, \dots, 11, \quad (40)$$

el mínimo valor que satisface estas restricciones debe ocurrir en algún borde de la región⁷.

por Jaynes¹¹, quien demostró que las soluciones de mayor entropía son inherentemente más probables, pues al tener mayor multiplicidad son las que pueden ser realizadas de más maneras por la naturaleza.

REFERENCIAS

1. C. Kittel, *"Introduction to Solid State Physics"*, John Wiley and sons, Inc, (1959).
2. D.H. Martin, *"Magnetism in Solids"*, Ilife Books Ltd., London, (1967).
3. *"Maximum Entropy Formalism"*, R. Levine y M. Tribus (editores), MIT, Boston, (1979).
4. A. Plastino, L. Rebollo Neira y A. Álvarez, "Statistical Inference, Signal Analysis and State Distribution", *Phys. Rev.*, Vol. 40, pp. 1644, (1989).
5. C. Shannon y W. Weaver, *"The Mathematical Theory of Communications"*, University of Illinois Press, Urbana, (1949).
6. E.T. Jaynes, "Information Theory and Statistical Mechanics", *Phys. Rev.*, Vol. 106, pp. 620, (1957).
7. G. Hadley, *"Nonlinear and Dynamic Programming"*, Addison-Wesley Publishing Co. Inc., Reading, Mass., (1964).
8. S. Brandt, *"Statistical and Computational Methods in Data Analysis"*, North Holland, Amsterdam, (1970).
9. D.W. Marquardt, *J. Soc. Indust. Appl. Math.*, Vol. 11, no. 2, (1963).
10. C.L. Lawson y R.J. Hanson, *"Solving Least Square Problems"*, Prentice Hall, Englewood Cliffs, N.J., (1974).
11. E.T. Jaynes, "Prior Information and Ambiguity in Inverse Problems", *SIAM-AMS Proceedings*, Vol. 14, pp. 151-166, (1984).