

Reconstruction d'un spectre RMN 2D par maximum d'entropie

Emilie CHOUZENOUX¹, Saïd MOUSSAOUI¹, Jérôme IDIER¹, François MARIETTE^{2,3}

¹IRCCYN, CNRS UMR 6597, 1, rue de la Noë, BP 92101, F-44321 Nantes Cedex 03, France
phone: + 33 2 40 37 69 24, fax: + 33 2 40 37 69 30
{emilie.chouzenoux, said.moussaoui, jerome.idier}@ircyn.ec-nantes.fr

²Cemagref, UR TERE CS 64427, 17 avenue de Cucillé, 35044 Rennes Cedex, France
phone: + 33 2 23 48 21 21, fax: + 33 2 23 48 21 15
francois.mariette@cemagref.fr

³Université Européenne de Bretagne, France

Thème – Problèmes inverses, restauration et reconstruction, bio-ingénierie et sciences de la vie

Problème traité – Estimation d'un spectre de corrélation $T_1 - T_2$ en résonance magnétique nucléaire 2D.

Originalité – Exploitation de la structure du problème de RMN 2D permettant d'éviter les approximations proposées récemment par Van-kataramanan *et coll.*. Proposition d'une méthode de reconstruction fondées sur le maximum d'entropie, et comparaison de deux algorithmes d'optimisation dédiés, l'un correspondant à une modification de l'algorithme classique dû à Skilling et Bryan ; l'autre à un algorithme de gradient conjugué non linéaire dont l'étape de recherche de pas tient compte du terme d'entropie.

Résultats – Application originale de la méthode du maximum d'entropie à la reconstruction 2D en RMN. Traitement de données synthétiques et réelles issues de l'analyse de produits agro-alimentaires.

1 Introduction

La spectroscopie par résonance magnétique nucléaire (RMN) est utilisée dans le domaine agro-alimentaire pour déterminer la composition (teneur en eau ou matière grasse) et accéder à des informations structurales aux échelles microscopique ou moléculaire. D'une façon générale, les mesures sont basées sur la réalisation d'enregistrements des temps de relaxation, T_1 et T_2 de façon indépendante. Des travaux ont démontrés que la mesure couplée de ces paramètres permettait un accroissement de la robustesse du calcul mais aussi une mise en évidence de couplages entre T_1 et T_2 aidant notablement à leur interprétation [1, 2]. Toutefois la réalisation de mesures couplées implique de disposer d'une méthode de traitement des spectres à 2 dimensions. Cet article présente une méthode basée sur le Maximum d'Entropie pour l'estimation d'un spectre de corrélation 2D. Le signal mesuré dépend à la fois de la durée t_1 correspondant au temps laissé au spin pour évoluer vers sa position d'équilibre (relaxation T_1) et de l'instant t_2 d'enregistrement de l'amplitude de l'écho RMN à t_1 fixé. Le modèle d'observation [2] relie ce signal RMN 2D, noté $Y(t_1, t_2)$, à la distribution des populations de molécules en fonction des temps de relaxation T_1 et T_2 , notée $S(T_1, T_2)$. Ces temps de relaxation caractérisent la dynamique des molécules en fonction de t_1 et t_2 respectivement. Ce modèle s'exprime par :

$$Y(t_1, t_2) = \iint k_1(t_1, T_1) S(T_1, T_2) k_2(t_2, T_2) dT_1 dT_2 + E(t_1, t_2) \quad (1)$$

avec des noyaux $k_1(t_1, T_1) = 1 - e^{-t_1/T_1}$ et $k_2(t_2, T_2) = e^{-t_2/T_2}$. Le bruit de mesure $E(t_1, t_2)$, est supposé i.i.d. gaussien centré. Après discrétisation, le modèle s'écrit sous la forme matricielle :

$$\mathbf{Y} = \mathbf{K}_1 \mathbf{S} \mathbf{K}_2^T + \mathbf{E} \quad \mathbf{K}_1 \in \mathbb{R}^{m_1 \times N_1}, \mathbf{K}_2 \in \mathbb{R}^{m_2 \times N_2}, \mathbf{Y} \in \mathbb{R}^{m_1 \times m_2}, \mathbf{S} \in \mathbb{R}^{N_1 \times N_2} \quad (2)$$

où m_1, m_2 représentent le nombre de points de mesure en t_1, t_2 et N_1, N_2 la taille de la grille T_1, T_2 .

L'objectif du traitement est d'estimer le spectre \mathbf{S} à partir de \mathbf{Y} en utilisant le modèle (2) et sous la contrainte $\mathbf{S} \geq 0$ (au sens où tous les éléments de \mathbf{S} sont non-négatifs). Ce problème de reconstruction est équivalent à une inversion numérique d'une

transformée de Laplace 2D, qui est très instable en raison du mauvais conditionnement des matrices \mathbf{K}_1 et \mathbf{K}_2 [2]. Afin d'obtenir une solution régularisée, les méthodes existantes minimisent classiquement un critère d'adéquation aux données $C(\mathbf{S})$ augmenté d'un terme de pénalisation $R(\mathbf{S})$, sous contrainte de positivité :

$$\min_{\mathbf{S} \geq 0} L(\mathbf{S}) = C(\mathbf{S}) + \beta R(\mathbf{S}). \quad (3)$$

Le bruit étant supposé gaussien, $C(\mathbf{S})$ est un critère de moindres carrés $C(\mathbf{S}) = \frac{1}{2} \|\mathbf{Y} - \mathbf{K}_1 \mathbf{S} \mathbf{K}_2^t\|_F^2$ où $\|\cdot\|_F$ représente la norme de Frobenius. Comme le spectre RMN 2D correspond à une distribution de populations de molécules, une pénalisation de type maximum d'entropie [3] nous paraît adaptée. Une méthode de reconstruction fondée sur le maximum d'entropie a d'ailleurs été précédemment proposée en RMN 1D [4].

Une autre difficulté concerne la dimension du problème. En effet, les méthodes existantes requièrent la réécriture du modèle direct sous la forme [2, 5] :

$$\mathbf{y} = \mathbf{K} \mathbf{s} + \mathbf{e}, \quad \mathbf{y} = \text{vect}(\mathbf{Y}), \mathbf{s} = \text{vect}(\mathbf{S}), \mathbf{e} = \text{vect}(\mathbf{E}), \mathbf{K} = \mathbf{K}_1 \otimes \mathbf{K}_2 \quad (4)$$

en notant vect la transformation matrice-vecteur dans l'ordre lexicographique et \otimes le produit de Kronecker. Néanmoins, ces méthodes ne sont plus applicables dès lors que la taille du problème augmente. En effet, une configuration typique en RMN est $m_1 = 50$, $m_2 = 10000$ et $N_1 \times N_2 = 200 \times 200$, ce qui conduit à une matrice \mathbf{K} de 2.10^{10} éléments. Pour contourner ce problème, [2] applique une décomposition SVD tronquée aux matrices \mathbf{K}_1 et \mathbf{K}_2 pour réduire la taille de la matrice \mathbf{K} . Or cette troncature est source de perte d'informations et peut être évitée. En effet, l'évaluation du gradient et du hessien du critère peut être réalisée avec une complexité réduite en tirant profit de la structure de la matrice \mathbf{K} , comme indiqué en fin de section 2.

2 Résolution du problème d'optimisation

Méthode de Bryan et Skilling. La méthode proposée par Skilling et Bryan [6] se fonde sur une approximation quadratique du critère dans un sous-espace \mathcal{D}_r de dimension r engendré à chaque itération k par $\mathbf{D}_k = [\mathbf{d}_k^{(1)} \dots \mathbf{d}_k^{(r)}]$. Les vecteurs $\mathbf{d}_k^{(j)}$ sont définis par $\mathbf{d}_k^{(1)} = \phi_k(\nabla L(\mathbf{s}_k))$ et par la récursion $\mathbf{d}_k^{(j+1)} = \phi_k(\nabla^2 L(\mathbf{s}_k) \mathbf{d}_k^{(j)})$, $j > 1$, ϕ_k étant une fonction d'échelle prenant en compte la positivité. L'équation de mise à jour de \mathbf{s} est alors donnée par :

$$\mathbf{s}_{k+1} = \mathbf{s}_k + \sum_{j=1}^r \xi_k^{(j)} \mathbf{d}_k^{(j)} = \mathbf{s}_k + \mathbf{D}_k \boldsymbol{\xi}_k. \quad (5)$$

Les $\xi_k^{(j)}$ résultent de la minimisation d'une approximation quadratique du critère $L(\mathbf{s})$ au point \mathbf{s}_k . Ils sont solutions de

$$(\mathbf{D}_k^t \nabla^2 L(\mathbf{s}_k) \mathbf{D}_k) \boldsymbol{\xi}_k = -\mathbf{D}_k^t \nabla L(\mathbf{s}_k). \quad (6)$$

Cet algorithme peut être vu comme un algorithme de descente $\mathbf{s}_{k+1} = \mathbf{s}_k + \alpha_k \boldsymbol{\delta}_k$ dans lequel la direction est $\boldsymbol{\delta}_k = \mathbf{D}_k \boldsymbol{\xi}_k$ et le pas α_k est fixé à 1. Cependant, le choix d'un pas unitaire ne garantit pas la décroissance du critère, ni même le respect de la contrainte de positivité. En pratique, si l'application de (5) fournit des composantes négatives, elles sont remplacées par une petite valeur positive ϵ , mais la décroissance du critère n'est pas non plus respectée à l'issue de cette opération.

La variante que nous proposons diffère de celle de Skilling et Bryan sous deux aspects. D'une part, un test portant sur la décroissance du critère est effectué, et une valeur de pas assurant la décroissance est adoptée dans la direction $\boldsymbol{\delta}_k$ si ce test est négatif ; la formule de pas utilisée est issue de travaux récents détaillés dans [7]. D'autre part, nous suggérons de sélectionner à chaque itération la valeur de r la plus favorable en terme de décroissance du critère, plutôt que de fixer r à une valeur constante. En effet, nous avons constaté que le comportement et l'efficacité de la méthode de Skilling et Bryan dépendait de façon critique du choix de r , et qu'il était difficile d'anticiper le meilleur choix. D'autre part, la stratégie que nous proposons n'entraîne pas d'augmentation du coût de calcul par itération, à condition de tirer parti de l'emboîtement des matrices normales des systèmes linéaires (6) à inverser quand r varie.

Méthode du gradient conjugué non-linéaire. D'après [8, p. 1022], la convergence mathématique de l'approche de Skilling et Bryan reste à prouver. Pour cette raison, nous proposons une autre famille d'algorithmes construits sur une base théorique plus solide, et plus spécifiquement un algorithme de gradient conjugué non-linéaire, utilisant une direction de descente donnée par :

$$\mathbf{d}_k = -\mathbf{c}_k \text{sign}(\mathbf{g}_k^t \mathbf{c}_k) \quad \text{avec} \quad \mathbf{c}_k = -\mathbf{g}_k + \beta_k \mathbf{d}_{k-1} \quad (7)$$

où β_k est le coefficient de conjugaison. La mise à jour $\mathbf{s}_{k+1} = \mathbf{s}_k + \alpha_k \mathbf{d}_k$ nécessite la recherche d'un pas α_k s'approchant du minimiseur de la fonction scalaire $\ell(\alpha) = L(\mathbf{s}_k + \alpha \mathbf{d}_k)$. Dans le cas du maximum d'entropie, $\ell(\alpha)$ (ou l'opposé de sa dérivée) tend

vers l'infini lorsque α se rapproche de $\bar{\alpha}$, plus petite valeur du pas pour laquelle une composante du vecteur $\mathbf{s}_k + \bar{\alpha}\mathbf{d}_k$ s'annule. Or la présence de cette barrière nuit à l'efficacité des méthodes classiques de recherche de pas, telles que le backtracking, la dichotomie et l'interpolation cubique. C'est pourquoi nous avons appliqué une stratégie de pas fondée sur la théorie des algorithmes Maximisation-Minimisation (MM) [9, 10]. Dans le cas d'une barrière logarithmique, la minimisation d'une fonction majorante de la forme :

$$h_k(\alpha, \alpha') = \ell(\alpha') + (\alpha - \alpha')\dot{\ell}(\alpha') + \frac{1}{2}m_k(\alpha - \alpha')^2 + \gamma_k \left[(\bar{\alpha} - \alpha') \log \left(\frac{\bar{\alpha} - \alpha'}{\bar{\alpha} - \alpha} \right) - \alpha + \alpha' \right] \quad (8)$$

est étudiée dans [7]. Les résultats assurent la convergence de l'algorithme GCNL vers la solution de (3) dans le cas de l'entropie de Burg, pour différentes formes de conjugaisons.

Mise en œuvre. Les algorithmes proposés peuvent être appliqués à la reconstruction d'un spectre RMN 2D sans la réduction de dimension suggérée par [2]. En effet, à l'aide d'opérations matricielles ne nécessitant pas la matrice \mathbf{K} , le gradient du critère d'adéquation aux données et le produit hessien-vecteur s'expriment par :

$$\nabla C(\mathbf{s}) = \text{vect}(-\mathbf{K}_1^t(\mathbf{Y} - \mathbf{K}_1\mathbf{S}\mathbf{K}_2^t)\mathbf{K}_2) \quad \text{et} \quad \nabla^2 C(\mathbf{s}) \cdot \mathbf{v} = \text{vect}(\mathbf{K}_1^t\mathbf{K}_1\mathbf{V}\mathbf{K}_2^t\mathbf{K}_2) \quad (9)$$

où \mathbf{V} est la transformation vecteur-matrice de \mathbf{v} .

3 Résultats expérimentaux

Des tests préliminaires sur des données synthétiques nous ont permis de valider les méthodes d'optimisation proposées. Nous avons ensuite appliqué ces deux méthodes pour estimer des spectres de corrélation $T_1 - T_2$ des molécules d'eau d'un échantillon de tissu végétal. Les mesures ont été effectuées pour $m_1 = 50$, $m_2 = 10000$ et la reconstruction a été réalisée pour $N_1 = 200$, $N_2 = 200$. L'algorithme de Bryan et Skilling est utilisé avec pour pénalisation l'entropie de Shannon. Pour l'algorithme GCNL, nous avons retenu l'entropie de Burg car les conditions de convergence énoncées dans [7] ne sont pas respectées dans le cas de l'entropie de Shannon. En choisissant correctement les paramètres de régularisation, les deux spectres obtenus sont très similaires.

L'analyse du comportement des algorithmes amène plusieurs conclusions. D'une part, la variante de l'algorithme de Skilling et Bryan est aussi rapide que la formulation classique dans sa version la plus rapide (c'est-à-dire optimisée empiriquement en fonction de r). De plus, elle permet d'assurer la décroissance du critère, donc un comportement plus régulier, ce qui faciliterait une analyse mathématique. En comparaison, l'algorithme GCNL est sensiblement plus lent, environ d'un facteur trois. Néanmoins, le potentiel d'accélération reste important, sans pour autant remettre en cause l'analyse théorique de convergence. En particulier, un algorithme de type L-BFGS serait utilisable d'après [7]. De plus, l'utilisation d'un préconditionneur exploitant la structure de la matrice $\mathbf{K} = \mathbf{K}_1 \otimes \mathbf{K}_2$ est sans doute un point crucial, étant donné le mauvais conditionnement du problème. Nous poursuivons actuellement nos travaux dans cette direction.

Références

- [1] A. E. English, K. P. Whittall, M. L. G. Joy et R. M. Henkelman, « Quantitative two-dimensional time correlation relaxometry », *Magn. Reson. Med.*, vol. 22, pp. 425–434, 1991.
- [2] Y. Q. Song, L. Venkataramanan, M. D. Hürlimann, M. Flaum, P. Frulla et C. Straley, « T1-T2 correlation spectra obtained using a fast two-dimensional Laplace inversion », *J. Magn. Reson.*, vol. 154, pp. 261–268, 2002.
- [3] A. Mohammad-Djafari et G. Demoment, « Maximum entropy image reconstruction in X-ray and diffraction tomography », *IEEE Trans. Medical Imaging*, vol. 7, n°4, pp. 345–354, 1988.
- [4] F. Mariette, J. P. Guillemet, C. Tellier et P. Marchal, « Continuous relaxation time distribution decomposition by MEM », *Signal Treat. and Signal Anal. in NMR*, pp. 218–234, 1996.
- [5] L. Venkataramanan, Y. Q. Song et M. D. Hürlimann, « Solving Fredholm integrals of the first kind with tensor product structure in 2 and 2.5 dimensions », *IEEE Trans. Signal Processing*, vol. 50, n°5, pp. 1017–1026, 2002.
- [6] J. Skilling et R. K. Bryan, « Maximum entropy image reconstruction : General algorithm », *Month. Not. Roy. Astr. Soc.*, vol. 211, pp. 111–124, 1984.
- [7] E. Chouzenoux, S. Moussaoui et J. Idier, « A new line search method for barrier functions with strong convergence properties », Rapport interne, IRCCyN, 2009, <http://hal.archives-ouvertes.fr/IRCCYN-ADTSI>.
- [8] W. H. Press, S. A. Teukolsky, W. T. Vetterling et B. P. Flannery, *Numerical Recipes : The Art of Scientific Computing*, Cambridge Univ. Press, New York, 3ème édition, 1992.
- [9] D. R. Hunter et K. Lange, « A tutorial on MM algorithms », *Amer. Statist.*, vol. 58, n°1, pp. 30–37, fév. 2004.
- [10] C. Labat et J. Idier, « Convergence of conjugate gradient methods with a closed-form stepsize formula », *J. Optim. Theory Appl.*, vol. 136, n°1, pp. 43–60, jan. 2008.