

Préconditionnement et régularisation des problèmes mal posés

A. Chabib, J.-F. Witz, V. Magnier, P. Gosselet

¹ Univ. Lille, CNRS, Centrale Lille, UMR 9013 - LaMcube - F-59000 Lille, France, {prenom.nom@univ-lille.fr}

Résumé — Ce travail s'intéresse au rôle du preconditionnement dans la régularisation des problèmes mal-posés résolus par gradient conjugué et son lien avec la méthode de Tikhonov. On montre qu'une analyse de Ritz permet d'estimer, à un coût de calcul négligeable, la solution pour n'importe quel poids de Tikhonov et de tracer des courbes en L et des graphes de Picard très utiles à l'analyse du système. Ce solveur linéaire amélioré est appliqué au problème de complétion de données de frontière, lors de la présentation on s'intéressera également à des problèmes de corrélation d'images.

Mots clés — problème inverse ; régularisation ; preconditionnement ; gradient conjugué.

1 Introduction

Les systèmes d'équations mal posés sont omniprésents en mécanique, particulièrement dans les problèmes d'identification, tels que la complétion de données de frontière en élasticité [8, 4]. Outre les questions d'existence et d'unicité, les problèmes mal posés se caractérisent par un manque de stabilité entre la cause et l'effet, en d'autres termes, de petites perturbations dans l'entrée peuvent entraîner d'importantes modifications de la sortie.

Dans cet article, nous nous concentrons sur les systèmes symétriques semi-définis positifs linéaires (linéarisés) discrets $n \times n$ de la forme $\mathbf{Ax} = \mathbf{b}$. Par exemple \mathbf{A} peut correspondre à la discrétisation d'un opérateur compact (voir les illustrations dans cet acte), où à la minimisation d'un critère aux moindres carrés (comme l'énergie image en corrélation d'images qui sera utilisée lors de la présentation). Les hypothèses permettent d'analyser toutes les propriétés en termes de spectre de \mathbf{A} , qui peut être diagonalisé comme $\mathbf{A} = \sum_{i=1}^n \sigma_i \mathbf{u}_i \mathbf{u}_i^T$ avec (\mathbf{u}_i) une base orthonormée et (σ_i) des valeurs propres non négatives par ordre décroissant.

L'existence et l'unicité sont liées au noyau de \mathbf{A} (valeurs propres strictement nulles) tandis que la stabilité est associée à l'accumulation de valeurs propres proches de zéro. En effet, une petite contribution de \mathbf{b} dans une direction propre associée à une petite valeur propre de \mathbf{A} a un impact significatif sur $\mathbf{x} = \sum \frac{\mathbf{u}_i^T \mathbf{b}}{\sigma_i} \mathbf{u}_i$. Les problèmes mal posés entraînent donc des opérateurs mal conditionnés.

Résoudre de tels systèmes revient à trouver un traitement satisfaisant pour ces petites valeurs propres : troncature, décalage, filtrage. Toutes ces techniques sont souvent contrôlées par un paramètre qui doit être ajusté afin de trouver un équilibre entre les informations contenues dans le système d'origine et celles apportées (ou supprimées) par le traitement. Lorsque la précision des données est connue, le principe de Morozov [9] fournit un critère objectif pour choisir le paramètre : la correction introduite par les informations ajoutées ne doit pas dépasser le bruit dans la mesure.

Lorsque ces données ne sont pas disponibles, il faut trouver un compromis. Le principe de Picard [5] compare les valeurs propres (σ_i) et la décomposition du second membre sur les directions propres $(\mathbf{u}_i^T \mathbf{b})$. Tant que les valeurs propres diminuent moins rapidement que les contributions du chargement, la solution reste contrôlée. La courbe en L [6] est une aide visuelle pour trouver un équilibre. Les solutions pour différents niveaux de régularisation sont positionnées dans un repère (« norme du résidu », « norme de la solution »). En général, une régularisation importante conduit à une norme faible de la solution mais à une erreur élevée, tandis qu'une régularisation faible conduit à un niveau d'erreur plus bas mais à des solutions importantes (fortement perturbées). Idéalement, il existe un point d'équilibre entre le résidu et la solution oscillante.

Résoudre un système mal posé à l'aide d'un solveur itératif peut sembler contre-intuitif, car un mauvais preconditionnement est souvent considéré comme un signal d'alarme pour l'utilisation de tels sol-

veurs. En réalité, cela peut être considéré comme une opportunité de mettre en œuvre approximativement les stratégies mentionnées ci-dessus : en raison de leur parenté avec les puissances itérées, les solveurs itératifs ont tendance à privilégier la partie supérieure du spectre au début et le nombre limité d'itérations (contrôlé par le seuil de convergence) peut être un moyen d'arrêter les itérations avant de chercher trop profondément dans la partie mal conditionnée du spectre. Cet article se concentre sur les nombreuses opportunités offertes par l'introduction d'un préconditionneur dans l'itération.

Le préconditionneur \mathbf{M} est souvent présenté dans les cours de méthodes numériques comme une approximation peu coûteuse à inverser de l'opérateur \mathbf{A} , dans le sens où le spectre de $\mathbf{M}^{-1}\mathbf{A}$ doit être aussi concentré que possible autour d'une valeur non nulle (qui peut être mise à l'échelle à 1), ce qui est grossièrement estimé par le conditionnement de $\mathbf{M}^{-1}\mathbf{A}$. Cependant, cela ne s'applique pas aux problèmes mal posés où l'inverse de \mathbf{A} ne peut pas être correctement défini. Quoiqu'il en soit, le préconditionneur peut jouer un rôle majeur, car il peut aider à atténuer la partie indésirable du spectre. Il est bien connu que le préconditionnement avec la matrice symétrique définie positive \mathbf{M} équivaut à résoudre le système $\mathbf{L}^{-1}\mathbf{A}\mathbf{L}^{-T}\mathbf{y} = \mathbf{L}^{-1}\mathbf{b}$ où $\mathbf{L}\mathbf{L}^T = \mathbf{M}$ est la factorisation de Cholesky, et $\mathbf{x} = \mathbf{L}^{-T}\mathbf{y}$. Ce système est régi par les valeurs propres de $\mathbf{L}^{-1}\mathbf{A}\mathbf{L}^{-T}$ ou, plus concrètement, par les valeurs propres généralisées (μ_i) de (\mathbf{A}, \mathbf{M}) : $\mathbf{A}\mathbf{v}_i = \mu_i\mathbf{M}\mathbf{v}_i$, avec (\mathbf{v}_i) une base \mathbf{M} -orthonormée. Les valeurs propres généralisées étant la valeur du quotient de Rayleigh $R(\mathbf{v}) = (\mathbf{v}^T\mathbf{A}\mathbf{v})/(\mathbf{v}^T\mathbf{M}\mathbf{v})$ aux points de stationnarité, on voit qu'en contrôlant le dénominateur, le préconditionneur peut servir à pénaliser les directions \mathbf{v} où $\mathbf{v}^T\mathbf{M}\mathbf{v}$ est grand et à retarder leur exploration par le solveur itératif. Nous voyons que le choix du préconditionneur obéit à des critères similaires à ceux du choix de la régularisation, d'où l'utilisation de la même matrice \mathbf{M} .

Cet article et la présentation associée montreront tous les bénéfices de ce choix. En particulier les possibilités d'analyse qui découlent à coût négligeable d'une résolution.

Notations Nous utilisons une police normale pour les scalaires, des minuscules en gras pour les vecteurs et des majuscules en gras pour les matrices. Une collection de vecteurs (\mathbf{x}_j) peut être représentée sous forme matricielle $\mathbf{X}_m = (\mathbf{x}_0, \dots, \mathbf{x}_{m-1})$, l'indice m correspondant alors au nombre de colonnes de la matrice. Ce travail est présenté dans \mathbb{R}^n , même si les méthodes s'appliquent également aux matrices hermitiennes et vecteurs complexes.

2 Gradient conjugué préconditionné, spectre de Ritz

Soit \mathbf{A} une matrice symétrique définie positive et \mathbf{b} un vecteur. On cherche la solution au problème $\mathbf{A}\mathbf{x} = \mathbf{b}$ en utilisant l'algorithme du gradient conjugué préconditionné par la matrice symétrique semi-définie positive \mathbf{M} .

À l'itération i , on note \mathbf{x}_i l'approximation et $\mathbf{r}_i = \mathbf{b} - \mathbf{A}\mathbf{x}_i$ le résidu. On introduit l'espace de Krylov $\mathcal{K}_i(\mathbf{M}^{-1}\mathbf{A}, \mathbf{M}^{-1}\mathbf{r}_0)$:

$$\mathcal{K}_i(\mathbf{M}^{-1}\mathbf{A}, \mathbf{M}^{-1}\mathbf{r}_0) = \text{span}\left(\mathbf{M}^{-1}\mathbf{r}_0, \dots, (\mathbf{M}^{-1}\mathbf{A})^{(i-1)}\mathbf{M}^{-1}\mathbf{r}_0\right) \quad (1)$$

Pour une initialisation arbitraire \mathbf{x}_0 , et le résidu associé $\mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0$, la i_{eme} itération est définie par l'équation suivante et mise en œuvre par l'algorithme 1.

$$\begin{cases} \text{trouver} & \mathbf{x}_i \in \mathbf{x}_0 + \mathcal{K}_i(\mathbf{M}^{-1}\mathbf{A}, \mathbf{M}^{-1}\mathbf{r}_0) \\ \text{tel que} & \mathbf{r}_i \perp \mathcal{K}_i(\mathbf{M}^{-1}\mathbf{A}, \mathbf{M}^{-1}\mathbf{r}_0) \end{cases} \quad (2)$$

L'algorithme construit deux bases spéciales de $\mathcal{K}_i(\mathbf{M}^{-1}\mathbf{A}, \mathbf{M}^{-1}\mathbf{r}_0)$, \mathbf{Z}_i est \mathbf{M} -orthogonale et \mathbf{W}_i est \mathbf{A} -orthogonale :

$$\begin{aligned} \mathbf{Z}_i^T \mathbf{M} \mathbf{Z}_i &= \mathbf{Z}_i^T \mathbf{R}_i = \text{diag}(\gamma_j)_{0 \leq j < i} \\ \mathbf{W}_i^T \mathbf{A} \mathbf{W}_i &= \mathbf{W}_i^T \mathbf{Q}_i = \text{diag}(\delta_j)_{0 \leq j < i} \end{aligned} \quad (3)$$

Il est pratique d'introduire la version \mathbf{M} -normalisée de la base \mathbf{Z}_i :

$$\hat{\mathbf{z}}_i = \frac{(-1)^i \mathbf{z}_i}{\sqrt{\gamma_i}} \quad \text{tel que} \quad \hat{\mathbf{Z}}_i^T \mathbf{M} \hat{\mathbf{Z}}_i = \mathbf{I}. \quad (4)$$

Algorithm 1 Gradient Conjugué préconditionné

```
 $\mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0 = \mathbf{P}^T \mathbf{r}_0$   
 $\mathbf{z}_0 = \mathbf{P}\mathbf{M}^{-1}\mathbf{r}_0, \mathbf{w}_0 = \mathbf{z}_0$   
 $\gamma_0 = (\mathbf{z}_0^T \mathbf{r}_0)$   
for  $i = 0, 1, \dots, m$  (convergence) do  
   $\mathbf{q}_i = \mathbf{A}\mathbf{w}_i$   
   $\delta_i = (\mathbf{w}_i^T \mathbf{q}_i), \alpha_i = \delta_i^{-1}\gamma_i$   
   $\mathbf{x}_{i+1} = \mathbf{x}_i + \mathbf{w}_i\alpha_i$   
   $\mathbf{r}_{i+1} = \mathbf{r}_i - \mathbf{q}_i\alpha_i$   
   $\mathbf{z}_{i+1} = \mathbf{P}\mathbf{M}^{-1}\mathbf{r}_{i+1}$   
   $\gamma_{i+1} = (\mathbf{z}_{i+1}^T \mathbf{r}_{i+1})$   
   $\beta_i = \gamma_i^{-1}\gamma_{i+1}$   
   $\mathbf{w}_{i+1} = \mathbf{z}_{i+1} + \mathbf{w}_i\beta_i$   
end for
```

$\hat{\mathbf{Z}}_i$ est en fait la base qui aurait été obtenue par la procédure d'Arnoldi [10], et on a :

$$\hat{\mathbf{Z}}_i^T \mathbf{A} \hat{\mathbf{Z}}_i = \mathbf{T}_i = \text{Tridiag}(\eta_{j-1}, \mu_j, \eta_j)$$
$$\text{avec } \mu_0 = \frac{1}{\alpha_0}, \quad \mu_j = \frac{1}{\alpha_j} + \frac{\beta_{j-1}}{\alpha_{j-1}}, \quad \eta_j = \frac{\sqrt{\beta_j}}{\alpha_j} \quad (5)$$

On peut diagonaliser $\mathbf{T}_i = \mathbf{\Xi}_i \mathbf{\Theta}_i \mathbf{\Xi}_i^T$ avec $\mathbf{\Theta}_i$ la matrice diagonale des valeurs propres triées dans l'ordre décroissant et $\mathbf{\Xi}_i$ la matrice orthonormale des vecteurs propres.

Les vecteurs de Ritz sont $\mathbf{V}_i = \hat{\mathbf{Z}}_i \mathbf{\Xi}_i$, et $\mathbf{\Theta}_i$ les valeurs de Ritz du système. Ils satisfont :

$$\mathbf{V}_i^T \mathbf{M} \mathbf{V}_i = \mathbf{I} \quad \text{and} \quad \mathbf{V}_i^T \mathbf{A} \mathbf{V}_i = \mathbf{\Theta}_i. \quad (6)$$

De manière à signifier la dépendance des valeurs et vecteurs de Ritz vis-à-vis de l'itération d'arrêt (m), ils sont notés avec un exposant (m) : $\mathbf{\Theta}_m = \text{diag}(\theta_j^{(m)})_{1 \leq j \leq i}$ et $\mathbf{V}_m = (\mathbf{v}_1^{(m)}, \dots, \mathbf{v}_i^{(m)})$. Quand le nombre d'itérations m augmente, les $(\theta_j^{(i)})_{1 \leq j \leq i}$ et $(\mathbf{v}_j^{(i)})_{1 \leq j \leq i}$ approchent les valeurs et vecteurs propres généralisés du couple (\mathbf{A}, \mathbf{M}) .

La convergence du gradient conjugué est directement liée à celle des valeurs de Ritz. Le préconditionneur a pour rôle d'améliorer le spectre du système pour faciliter la convergence. Il faut noter que de par la proximité avec les méthodes de puissances itérées, la partie haute de spectre est explorée préférentiellement.

2.1 Normes naturelles de l'erreur et de la solution

Le préconditionneur définit des mesures de l'erreur et de la solution auxquelles on a accès au cours des itérations. On a les estimateurs d'erreur classiques [7, 1] :

$$\|\mathbf{r}_i\|_{\mathbf{M}^{-1}}^2 = \gamma_i$$
$$\|\mathbf{x}_{i+1} - \mathbf{x}\|_{\mathbf{A}}^2 = \|\mathbf{x}_i - \mathbf{x}\|_{\mathbf{A}}^2 - \gamma_i^2 \delta_i \quad (7)$$

On a également accès à la correction de la solution [4] :

$$\|\mathbf{x}_{i+1} - \mathbf{x}_0\|_{\mathbf{M}}^2 = \|\mathbf{x}_i - \mathbf{x}_0\|_{\mathbf{M}}^2 + \alpha_i^2 \|\mathbf{w}_i\|_{\mathbf{M}}^2 + 2\alpha_i (\mathbf{w}_i^T \mathbf{M} (\mathbf{x}_i - \mathbf{x}_0))$$
$$\text{avec } \begin{cases} \|\mathbf{w}_{i+1}\|_{\mathbf{M}}^2 = \gamma_i + \beta_i^2 \|\mathbf{w}_i\|_{\mathbf{M}}^2, & \|\mathbf{w}_0\|_{\mathbf{M}}^2 = \gamma_0, \\ (\mathbf{w}_{i+1}^T \mathbf{M} (\mathbf{x}_{i+1} - \mathbf{x}_0)) = -\beta_i ((\mathbf{w}_i^T \mathbf{M} (\mathbf{x}_i - \mathbf{x}_0)) + \alpha_i \|\mathbf{w}_i\|_{\mathbf{M}}^2) > 0. \end{cases} \quad (8)$$

Et enfin, on a accès à une estimation de l'opérateur :

$$\|\mathbf{T}_0\|_F^2 = \mu_0^2,$$
$$\|\mathbf{T}_{i+1}\|_F^2 = \|\mathbf{T}_i\|_F^2 + \mu_i^2 + \eta_i^2 + \eta_{i-1}^2 \rightarrow \|\mathbf{M}^{-\frac{1}{2}} \mathbf{A} \mathbf{M}^{-\frac{1}{2}}\|_F^2, \quad (9)$$

où F est la norme de Frobenius, $\|\mathbf{M}^{-\frac{1}{2}}\mathbf{A}\mathbf{M}^{-\frac{1}{2}}\|_F^2$ est la somme des carrés des valeurs propres généralisées du couple (\mathbf{A}, \mathbf{M}) . De là, on peut tirer le critère d'arrêt suivant :

$$\|\mathbf{r}_i\|_{\mathbf{M}^{-1}} < \varepsilon \|\mathbf{T}_i\|_F \|\mathbf{x}_i - \mathbf{x}_0\|_{\mathbf{M}}, \quad (10)$$

qui a l'avantage de mettre en regard l'évolution de la solution et celle de l'erreur.

Par ailleurs, $(\|\mathbf{x}_i - \mathbf{x}_0\|_{\mathbf{M}}^2, \|\mathbf{x}_{i+1} - \mathbf{x}\|_{\mathbf{A}}^2)$ forme un repère très intéressant pour tracer une courbe en L : par construction, l'évolution $i \rightarrow (i+1)$ se fait toujours vers le haut et la gauche.

2.2 Filtrage a posteriori

Les éléments de Ritz offrent un moyen pratique de filtrer la solution. En supposant que m itérations aient été effectuées, nous pouvons traiter la base \mathbf{V}_m et les valeurs Θ_m . Nous pouvons décomposer le second membre sur la base de Ritz $r_j^{(m)} = \mathbf{v}_j^{(m)T} \mathbf{r}_0$, et définir :

$$\text{pour } i \leq m, \quad \tilde{\mathbf{x}}_i^{(m)} = \mathbf{x}_0 + \sum_{j=1}^i \frac{r_j^{(m)}}{\theta_j^{(m)}} \mathbf{v}_j^{(m)}. \quad (11)$$

On a :

$$\begin{aligned} \|\tilde{\mathbf{x}}_i^{(m)} - \mathbf{x}\|_{\mathbf{A}}^2 &= \|\tilde{\mathbf{x}}_i^{(m)} - \mathbf{x}_0\|_{\mathbf{A}}^2 - \sum_{j=1}^i \frac{(r_j^{(m)})^2}{\theta_j^{(m)}}, \\ \|\tilde{\mathbf{x}}_i^{(m)} - \mathbf{x}_0\|_{\mathbf{M}}^2 &= \sum_{j=1}^i \frac{(r_j^{(m)})^2}{(\theta_j^{(m)})^2}, \end{aligned} \quad (12)$$

et, bien sûr :

$$\|\tilde{\mathbf{x}}_i^{(m)} - \tilde{\mathbf{x}}_{i-1}^{(m)}\|_{\mathbf{A}}^2 = \frac{(r_i^{(m)})^2}{\theta_i^{(m)}} \quad \text{et} \quad \|\tilde{\mathbf{x}}_i^{(m)} - \tilde{\mathbf{x}}_{i-1}^{(m)}\|_{\mathbf{M}}^2 = \frac{(r_i^{(m)})^2}{(\theta_i^{(m)})^2}. \quad (13)$$

Comme les $(\theta_i^{(m)})$ sont triés par ordre décroissant, nous voyons que l'erreur de $i \mapsto (\tilde{\mathbf{x}}_i^{(m)})$ tend à diminuer plus lentement que sa norme tend à augmenter. La courbe en L pour $(\tilde{\mathbf{x}}_i^{(m)})_i$ est alors convexe et ne zigzague pas.

La pente de la courbe en L entre le point $i-1$ et i est $-(\theta_i^{(m)})^{-1}$. Une possibilité consiste à définir le coin comme le point qui maximise la variation de la pente : $i = \arg \max_j ((\theta_{j+1}^{(m)})^{-1} - (\theta_j^{(m)})^{-1})$.

Les éléments de Ritz permettent également d'utiliser la théorie de Picard et d'arrêter la construction de $\tilde{\mathbf{x}}_i^{(m)}$ lorsque la contribution $j \mapsto r_j^{(m)}$ commence à diminuer moins rapidement que $j \mapsto \theta_j^{(m)}$. Ce critère a l'avantage de prendre en compte les propriétés du second membre.

3 Systèmes régularisés par Tikhonov

On s'intéresse aux systèmes où la matrice est régularisée par le preconditionneur : $\mathbf{A}_\lambda = \mathbf{A} + \lambda \mathbf{M}$, en supposant naturellement que résoudre un système sur \mathbf{M} peut être effectué à coût modéré. On considère que le second membre admet une dépendance affine vis-à-vis de λ : $\mathbf{b}_\lambda = \mathbf{b}_A + \lambda \mathbf{b}_M$ mais des dépendances séparées plus complexes seraient possibles ($\mathbf{b}(\lambda) = \sum_a f_a(\lambda) \mathbf{b}_a$) :

$$\underbrace{(\mathbf{A} + \lambda \mathbf{M})}_{\mathbf{A}_\lambda} \mathbf{x}_\lambda = \underbrace{\mathbf{b}_A + \lambda \mathbf{b}_M}_{\mathbf{b}_\lambda}. \quad (14)$$

On suppose que le système (14) a été résolu pour un λ donné en m itérations en utilisant le gradient conjugué preconditionné par \mathbf{M} . On peut alors calculer la base de Ritz \mathbf{V}_m . Le point fort est que \mathbf{V}_m sépare les effets de l'opérateur \mathbf{A} et ceux de la régularisation \mathbf{M} indépendamment de λ :

$$\begin{aligned} \mathbf{V}_m^T \mathbf{M} \mathbf{V}_m &= \mathbf{I}_m, \\ \mathbf{V}_m^T \mathbf{A}_\lambda \mathbf{V}_m &= \Theta_{\lambda, m} = \Theta_m + \lambda \mathbf{I}_m. \end{aligned} \quad (15)$$

λ peut être vu comme un décalage dans le spectre généralisé de (\mathbf{A}, \mathbf{M}) .

Le résidu initial prend la forme séparée :

$$\mathbf{r}_{\lambda,0} = \mathbf{b}_\lambda - \mathbf{A}_\lambda \mathbf{x}_0 = \underbrace{(\mathbf{b}_\mathbf{A} - \mathbf{A} \mathbf{x}_0)}_{\mathbf{r}_{\mathbf{A},0}} + \lambda \underbrace{(\mathbf{b}_\mathbf{M} - \mathbf{M} \mathbf{x}_0)}_{\mathbf{r}_{\mathbf{M},0}}, \quad (16)$$

et on peut définir les contributions spectrales $r_{\mathbf{A},j}^{(m)} = \mathbf{v}_j^{(m)T} \mathbf{r}_{\mathbf{A},0}$ et $r_{\mathbf{M},j}^{(m)} = \mathbf{v}_j^{(m)T} \mathbf{r}_{\mathbf{M},0}$,

Après m itérations, on peut définir l'approximation de Ritz pour $i \leq m$:

$$\tilde{\mathbf{x}}_{\lambda,i}^{(m)} = \mathbf{x}_0 + \sum_{j=1}^i \mathbf{v}_j^{(m)} \frac{(\mathbf{v}_j^{(m)T} \mathbf{r}_{\lambda,0})}{\theta_j^{(m)} + \lambda} = \mathbf{x}_0 + \sum_{j=1}^i \frac{r_{\mathbf{A},j}^{(m)} + \lambda r_{\mathbf{M},j}^{(m)}}{\theta_j^{(m)} + \lambda} \mathbf{v}_j^{(m)} \quad (17)$$

Ces approximations peuvent être calculées à un coût marginal, et leur dépendance en λ est explicite : la courbe en L de $\lambda \mapsto \tilde{\mathbf{x}}_{\lambda,i}^{(m)}$ est une fraction rationnelle. Elle permet même de donner un sens à la solution limite lorsque $\lambda \rightarrow 0$, même lorsque \mathbf{A} n'était pas inversible. Elle fournit également une formule analytique pour la recherche du choix optimal de (λ, i) réalisant un bon compromis entre l'erreur et la norme de la solution. Nous avons les propriétés suivantes :

$$\begin{aligned} \|\tilde{\mathbf{x}}_{\lambda,i}^{(m)} - \mathbf{x}_0\|_{\mathbf{M}}^2 &= \sum_{j=1}^i \left(\frac{r_{\mathbf{A},j}^{(m)} + \lambda r_{\mathbf{M},j}^{(m)}}{\theta_j^{(m)} + \lambda} \right)^2 \\ \|\tilde{\mathbf{x}}_{\lambda,i}^{(m)} - \mathbf{x}\|_{\mathbf{A}}^2 - \|\mathbf{x}_0 - \mathbf{x}\|_{\mathbf{A}}^2 &= \sum_{j=1}^i \frac{(r_{\mathbf{A},j}^{(m)} + \lambda r_{\mathbf{M},j}^{(m)})}{(\theta_j^{(m)} + \lambda)} \left(\frac{\theta_j^{(m)} (r_{\mathbf{A},j}^{(m)} + \lambda r_{\mathbf{M},j}^{(m)})}{(\theta_j^{(m)} + \lambda)} - 2r_{\mathbf{A},j}^{(m)} \right) \end{aligned} \quad (18)$$

4 Validation

Nous utilisons l'illustration classique du problème mal posé pour le problème inverse de Laplace. Nous considérons le domaine rectangulaire $\Omega := [0, T] \times [0, H]$, où le problème de Cauchy suivant est donné :

$$\begin{aligned} \Delta u &= 0 && \text{in } \Omega, \\ u &= 0 && \text{on } y = 0 \text{ et } y = H, \\ \frac{\partial u}{\partial x} &= 0 && \text{on } x = 0, \\ u &= u_L := \sin k\pi \frac{y}{H} && \text{on } x = 0, \end{aligned} \quad (19)$$

où $k \in \mathbb{N}$ est le nombre d'onde du signal. Comme on peut le constater, il n'y a pas de condition aux limites du côté droit $\Gamma_R = \{(T, y), y \in (0, H)\}$, alors qu'il existe à la fois des conditions de Dirichlet et de Neumann du côté gauche. Il s'agit d'un problème modèle pour le contrôle non destructif des structures en statique, qui admet la solution suivante :

$$u(x, y) = \sin k\pi \frac{y}{H} \cosh(k\pi \frac{x}{H}). \quad (20)$$

On observe que la solution du côté droit ($x = T$) explose pour les domaines épais ($T \nearrow$) ou les grands nombres d'onde ($k \nearrow$), comme l'illustre la figure 1. Ce type de problème est souvent qualifié de « sévèrement mal posé » [2].

On propose de résoudre ce système par la méthode de Steklov-Poincaré [3]. La place manque pour la décrire. On dira simplement que le problème est posé en termes de déplacement sur le bord droit $x = u_R$ que l'on cherche de manière que les deux problèmes bien posés avec les conditions de Dirichlet ou les conditions de Neumann connues à gauche conduisent à la même solution. On se retrouve avec un système de la forme $\mathbf{A} \mathbf{x} = \mathbf{b}$ qui est la discrétisation d'un problème compact. Même si cela est rarement présenté ainsi, la méthode classique pour résoudre ce système, dite de KMF, peut s'interpréter comme un point fixe couplé à un préconditionnement par un opérateur \mathbf{M} qui mesure l'énergie de la solution. On propose d'utiliser $\lambda \mathbf{M}$ pour régulariser le système. Pour information, dans cette configuration, les cinq

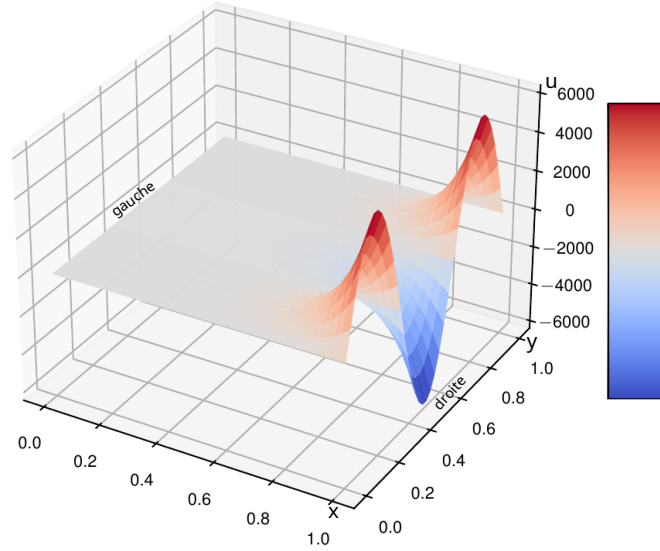
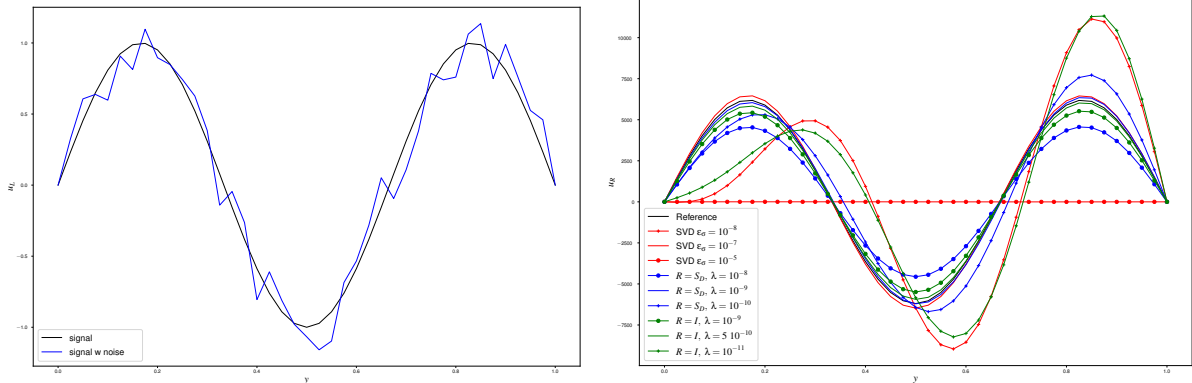


FIGURE 1 – Solution à identifier en observant uniquement le côté gauche (oscillation indiscernable), pour $k = 3$ et $H = T = 1$.

plus grandes valeurs propres de \mathbf{A} sont $\{5.8 \cdot 10^{-4}, 2.1 \cdot 10^{-6}, 5.6 \cdot 10^{-9}, 1.2 \cdot 10^{-11}, 2.3 \cdot 10^{-14}\}$ alors que le spectre de \mathbf{M} appartient à l'intervalle $[7.8 \cdot 10^{-3}, 1.63]$.

La figure 2b présente les champs identifiés en utilisant soit une SVD tronquée où seules les valeurs plus grandes que $\varepsilon_\sigma \sigma_1$ sont utilisées, ou l'opérateur régularisé $(\mathbf{A} + \lambda \mathbf{R})$, \mathbf{R} étant soit l'identité \mathbf{I} soit \mathbf{M} .



(a) Signal d'entrée u_L (bruit 10 dB)

(b) Champ identifié u_R avec SVD tronquée ou régularisation de Tikhonov

La figure 3 présente les résultats obtenus par des solveurs itératifs en fonction du préconditionneur et de la régularisation ($\mathbf{R} = \mathbf{M}$). Pour comparer avec le préconditionneur recommandé (KMF \mathbf{M}), on a utilisé le préconditionneur diagonal de Jacobi qui, s'il est rarement suffisant, ne fait pas de mal sur les problèmes bien posés. On voit qu'ici il est dramatique, sauf si l'opérateur est régularisé. On observe également l'intérêt d'utiliser les normes naturelles du gradient conjugué pour tracer des courbes en L (en fonction de l'itération) dont la forme est exploitable.

On considère maintenant l'opérateur de KMF utilisé à la fois comme régularisation et comme préconditionneur : $(\mathbf{A} + \lambda \mathbf{M})\mathbf{x} = \mathbf{b}$. Après les m itérations du gradient conjugué nécessaires pour converger, on post-processe les valeurs et les vecteurs de Ritz $\Theta^{(m)}$, $\mathbf{V}^{(m)}$.

La figure 4 présente l'approximation de la courbe en L en fonction de λ donnée par les éléments de Ritz post-traités (18) à partir d'un calcul itératif avec $\varepsilon = 10^{-9}$ et $\lambda = 10^{-9}$. Elle est comparée à la courbe en L obtenue à partir de résolutions directes avec divers $\lambda \in [10^{-6}, 10^{-12}]$, en utilisant le fait que dans ce cas théorique, la référence \mathbf{x} est connue. Notez que, comme pour le solveur itératif, l'erreur réelle est connue au facteur additif $\|x - x_0\|_{\mathbf{A}}^2$ près, les courbes sont en fait translatées de manière que leur asymptote verticale soit alignée avec l'axe y .

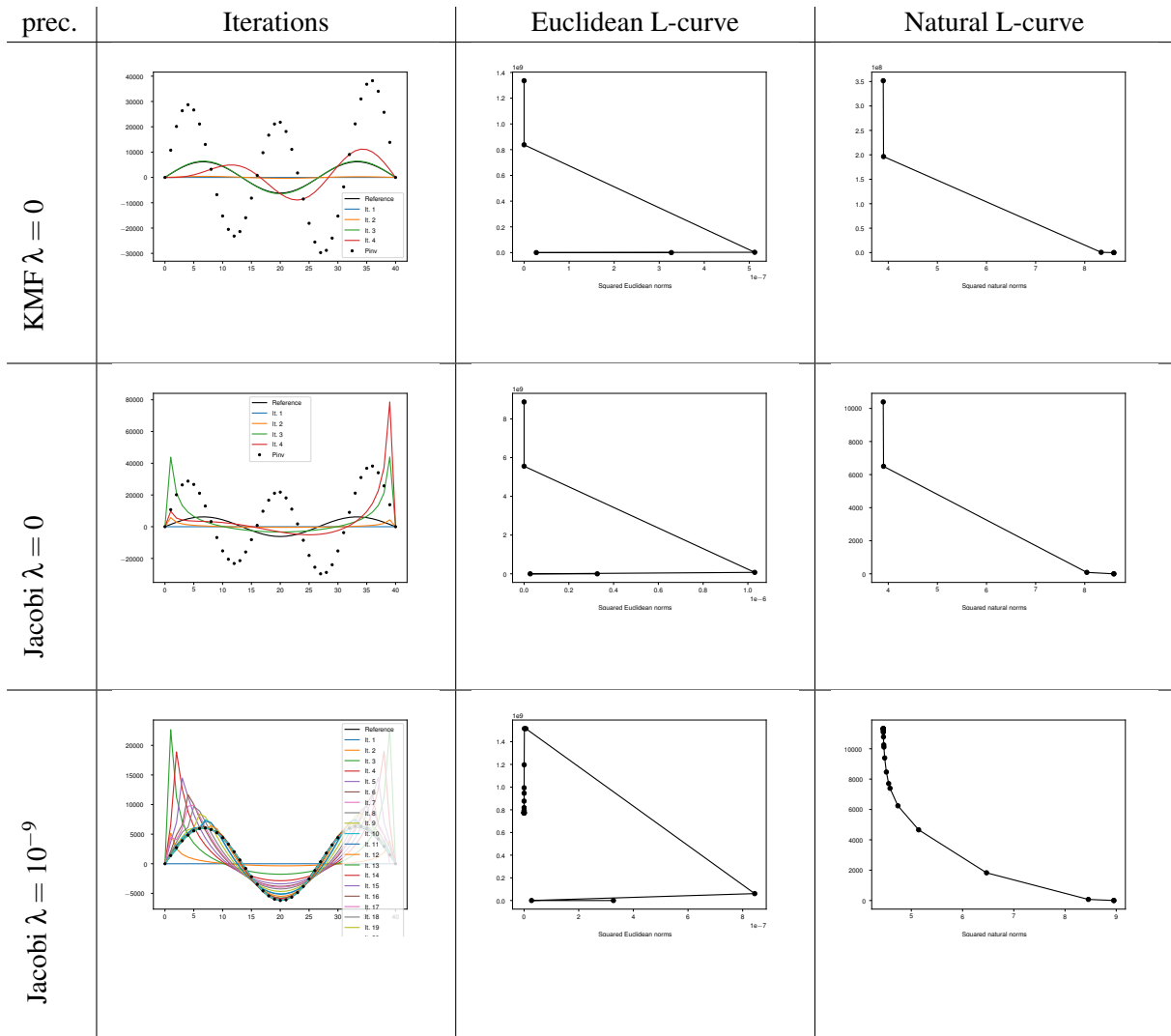


FIGURE 3 – Problème de complétion de données : qualité de la solution du gradient conjugué pour différents préconditionneurs – effet d’une légère régularisation sur le moins bon préconditionneur.

Nous voyons que la courbe en L post-traitée à un coût négligeable après la solution itérative fournit une excellente estimation de la courbe en L obtenue par résolution directe d’une famille de systèmes régularisés pour différents poids λ .

La figure 5 présente le graphique de Picard obtenu après la résolution itérative avec $\lambda = 10^{-9}$ et $\varepsilon = 10^{-12}$ (pour ce cas simple, une plus grande précision était nécessaire afin d’avoir plus de points à tracer). Tout d’abord, nous pouvons observer le spectre de \mathbf{A} estimé par les valeurs de Ritz. Grâce à (15), la régularisation est une ligne horizontale, et nous pouvons voir directement comment la régularisation influe sur le spectre : dans ce cas, les deux plus petites valeurs propres sont rectifiées.

Un deuxième intérêt du graphique de Picard est de comparer le spectre avec la contribution du second membre $|r_j^{(m)}|$. Il est clair qu’à part pour les deux premières composantes, les contributions du second membre diminuent plus lentement que les valeurs de Ritz. Cela indique que le système est instable, car l’effet des contributions à la solution devient de plus en plus important.

Ces deux informations observables dans les graphiques de Picard permettent de choisir en connaissance de cause λ et l’indice de troncature i dans la reconstruction de Ritz $\tilde{\mathbf{x}}_{\lambda,i}^{(m)}$.

5 Conclusion

Ce papier montre que la régularisation de Tikhonov et le préconditionnement obéissent à des logiques similaires quant au traitement du spectre des problèmes mal-posés. Quand elle est raisonnablement facile à résoudre, une bonne régularisation fournit donc un excellent préconditionneur qui permet de chercher

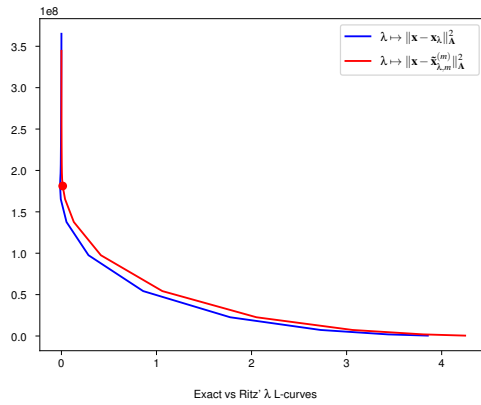


FIGURE 4 – L-courbe obtenue avec un solveur direct vs L-courbe estimée par l’approximation de Ritz (18) après une résolution itérative avec $\lambda = 10^{-9}$ matérialisée pour le point ($\varepsilon = 10^{-9}$). Les asymptotes verticales sont alignées avec l’axe des ordonnées.

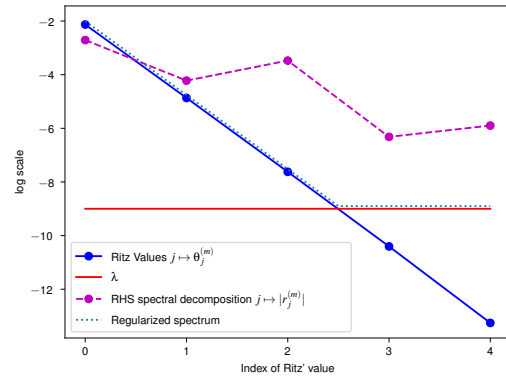


FIGURE 5 – Graphe de Picard obtenu après la résolution itérative avec $\lambda = 10^{-9}$ et $\varepsilon = 10^{-12}$.

naturellement des solutions régulières que les itérations viennent enrichir avec des détails. En outre quand la même matrice est choisie aux deux postes, une analyse de Ritz permet d’analyser très finement le système par des courbes de Picard et des courbes en L et de post-processor à coût négligeable l’ensemble des solutions pour divers poids de régularisation.

Références

- [1] Owe Axelsson and Igor Kaporin. Error norm estimation and stopping criteria in preconditioned conjugate gradient iterations. *Numerical Linear Algebra with Applications*, 8(4) :265–286, 2001.
- [2] Faker Ben Belgacem. Why is the Cauchy problem severely ill-posed? *Inverse Problems*, 23(2) :823–836, 2007.
- [3] Faker Ben Belgacem and Henda El Fekih. On Cauchy’s problem : I. A variational Steklov–Poincaré theory. *Inverse Problems*, 21(6) :1915, 2005.
- [4] Renaud Ferrier, Mohamed L. Kadri, and Pierre Gosselet. The Steklov-Poincaré technique for data completion : Preconditioning and filtering. *International Journal for Numerical Methods in Engineering*, 116(4) :270–286, 2018.
- [5] Per Christian Hansen. The discrete Picard condition for discrete ill-posed problems. *BIT Numerical Mathematics*, 30(4) :658–672, 1990.
- [6] Per Christian Hansen. Analysis of discrete ill-posed problems by means of the L-curve. *SIAM review*, 34(4) :561–580, 1992.
- [7] Magnus R. Hestenes and Eduard Stiefel. Methods of conjugate gradients for solving linear systems. *Journal of research of the national bureau of standards*, 49(6) :409–436, 1952.
- [8] Mohamed Larbi Kadri, Jalel Ben Abdallah, and Thouraya Nouri Baranger. Identification of internal cracks in a three-dimensional solid body via Steklov–Poincaré approaches. *Comptes Rendus Mécanique*, 339(10) :674–681, 2011.
- [9] Vladimir Alekseevich Morozov. The error principle in the solution of operational equations by the regularization method. *Zhurnal Vychislitel’noi Matematiki i Matematicheskoi Fiziki*, 8(2) :295–309, 1968.
- [10] Yousef Saad. *Iterative methods for sparse linear systems*. SIAM, 2003.