
Ecole doctorale Peyresq

Quelques aspects mathématiques et numériques des problèmes inverses linéaires

`pierre.armand.weiss@gmail.com`

Table des matières

1	Préliminaires	4
1.1	Avertissement	4
1.2	Un peu de théorie de la mesure	4
1.3	Quelques espaces fonctionnels	5
2	Quelques exemples de problèmes inverses linéaires	8
2.1	La déconvolution	8
2.2	Défloutage vs déconvolution	12
2.3	Super-résolution en microscopie optique/SMLM	13
2.4	L'échantillonnage compressé	15
2.5	Le cadre de travail général	15
3	Opérateurs intégraux	18
3.1	Définitions	18
3.2	Effet régularisant	20
3.3	Opérateurs de Hilbert-Schmidt	21
3.4	Théorie spectrale	23
3.5	Discrétisation d'opérateurs	24
3.6	Calcul matriciel rapide	27
4	Rappels d'algèbre linéaire	42
4.1	Image et noyau d'une matrice	42
4.2	Déterminant	42
4.3	Valeurs propres et vecteurs propres	43
4.4	Produit hermitien et produit scalaire	43
4.5	Isométries, matrices unitaires, matrices orthonormales et orthogonales	44
4.6	Normes matricielles	45
4.7	Valeurs singulières et valeurs propres	46

1 Préliminaires

1.1 Avertissement

Ces notes de cours ont été rédigées très rapidement pour me forcer à préparer ce cours doctoral et à produire un contenu que j'espère enrichissant pour certains. Elles sont probablement bourrées de coquilles et je m'en excuse. N'hésitez pas à me les indiquer à l'adresse pierre.armand.weiss@gmail.com. Je pourrais ainsi les corriger.

De plus, je vois ces notes plus comme un pense-bête que comme une référence. Je survole la majorité des concepts alors qu'il faudrait y passer bien plus de temps pour permettre de les intégrer plus profondément. Mon objectif en 5 heures est plutôt de présenter quelques concepts qui me semblent utiles pour la co-conception et d'essayer de démystifier quelques aspects pour vous permettre de les aborder plus sereinement.

Je vous encourage donc vivement à explorer les éléments bibliographiques indiqués pour approfondir et préciser les points qui vous semblent les plus pertinents. Si les planètes s'alignent, je m'en servirais peut-être comme une base pour un projet plus ambitieux de livre.

1.2 Un peu de théorie de la mesure

Dans tout le cours, le symbole \mathbb{K} pourra représenter l'ensemble des nombres réels \mathbb{R} ou l'ensemble des nombres complexes \mathbb{C} .

Dans tout ce cours, les images seront définies sur un espace mesuré (S, \mathcal{E}, ν) . Ici, S représente le domaine de définition d'une image, \mathcal{E} est une tribu sur S et ν une mesure. Le cadre ci-dessus peut paraître abstrait ou abscons suivant sa formation. Dans notre cours, il permettra notamment de traiter les problèmes discrets et continus sous une forme unifiée. De plus, on verra que certains problèmes s'expriment naturellement dans ce cadre.

Dans le cas discret, on posera simplement $S = \{1, \dots, N\}$ où N est le nombre de pixels d'une image, ν est la mesure de comptage et \mathcal{E} est l'ensemble des sous-ensembles de S , souvent noté $\mathcal{P}(S)$.

Dans ce cas continu on s'intéressera uniquement au cadre suivant. Une image sera définie sur un sous-domaine S de \mathbb{R}^D ou du tore \mathbb{T}^D (la raison principale est qu'on peut alors travailler avec des séries de Fourier). Ainsi, on posera $S \subset \mathbb{R}^D$, \mathcal{E} est la tribu des boréliens et ν est la mesure de Lebesgue sur S . On préférera souvent la notation ds à la notation $d\nu(s)$ pour définir des intégrales.

Je rappelle ci-dessous les définitions essentielles de la théorie de la mesure.

Définition 1.1 (Tribu). Une tribu (σ -algèbre) \mathcal{E} est un sous ensemble de $\mathcal{P}(S)$ telle que

- $\omega \in \mathcal{E} \Rightarrow \omega^c \in \mathcal{E}$.
- $S \in \mathcal{E}$
- $\bigcup_{i=1}^{\infty} \omega_i \in \mathcal{E}$ pour tout $(\omega_1, \omega_2, \dots) \in \mathcal{E}$.

Les conditions précédentes impliquent notamment que $\emptyset \in \mathcal{E}$ et que $\bigcap_{i=1}^{\infty} \omega_i \in \mathcal{E}$ pour tout $(\omega_1, \omega_2, \dots) \in \mathcal{E}$. Etant donné un espace S et sa tribu \mathcal{E} , on peut définir une mesure.

Définition 1.2 (Mesure sur \mathcal{E}). Une mesure μ sur un espace (S, \mathcal{E}) est une fonction sur \mathcal{E} telle que

- $\mu(\omega) \geq 0, \forall \omega \in \mathcal{E}$.
- $\mu(\emptyset) = 0$.
- $\mu(\cup_{i=1}^{\infty} \omega_i) = \sum_{i=1}^{\infty} \mu(\omega_i)$ pour tout ensemble d'ensembles $(\omega_1, \omega_2, \dots)$ dis-joints.

Si de plus $\mu(S) = 1$, μ est une mesure de probabilité.

Par définition, les ensembles mesurables sont les éléments de la tribu \mathcal{E} . Qu'est-ce qu'une fonction mesurable?

Définition 1.3 (Fonction mesurable). Une fonction mesurable $f : S \rightarrow \mathbb{K}$ est une fonction qui satisfait $f^{-1}(I) \in \mathcal{E}$ pour tout borélien I .

C'est cette définition qui permet de définir proprement l'intégrale de Lebesgue qu'on utilisera largement dans ce cours. Nous supposons cette définition acquise.

1.3 Quelques espaces fonctionnels

Nous verrons les images comme des éléments d'un espace vectoriel \mathcal{X} de fonctions mesurables définies sur S , muni d'une norme $\|\cdot\|_{\mathcal{X}}$. En pratique, seuls les choix suivant nous intéresseront.

Le cadre discret Le cadre le plus important (car il correspond à la grande majorité des travaux en traitement du signal et à la pratique numérique) est simplement :

$$\mathcal{X} = \mathbb{R}^N$$

où N est un nombre de pixels. Cet espace vectoriel peut être muni des normes ℓ^p standard notées $\|\cdot\|_p$.

De façon plus générale, on peut aussi se donner une famille de fonctions (e_1, \dots, e_N) d'un domaine continu S dans \mathbb{R} et définir les images comme des éléments du sous-espace vectoriel

$$\mathcal{X} = \left\{ x = \sum_{n=1}^N \alpha_n e_n, \alpha \in \mathbb{R}^N \right\} = \text{span}(e_1, \dots, e_N).$$

Le cadre L^2 Etant donné un espace mesuré (S, \mathcal{E}, ν) on pose

$$\mathcal{X} = L^2(S) \stackrel{\text{def}}{=} \left\{ x : S \rightarrow \mathbb{R} \text{ mesurables, } \int_S x^2(s) ds < +\infty \right\}.$$

On munit \mathcal{X} du produit scalaire usuel défini pour tout $x, x' \in \mathcal{X}$ par

$$\langle x, x' \rangle \stackrel{\text{def}}{=} \int_S x(s)x'(s) ds$$

et on définit la norme sur \mathcal{X} par $\|x\|_{L^2(S)} = \sqrt{\langle x, x \rangle}$.

C'est un cadre très commun en traitement du signal et en optimisation qui permet d'introduire les bases Hilbertiennes. Dans ce cadre, on peut en effet construire une famille de fonctions (e_1, e_2, \dots) orthogonales :

$$\langle e_n, e'_n \rangle = 0 \text{ si } n \neq n', \quad \text{et} \quad \langle e_n, e'_n \rangle = 1 \text{ sinon.}$$

telle que

$$\forall x \in \mathcal{X}, x = \sum_{n=0}^{\infty} \langle x, e_n \rangle e_n.$$

Voici quelques exemples :

- Sur $S = \mathbb{T}^D$ ou $S = [0, 1]^D$, on peut considérer les atomes de Fourier définis par $e_n(s) = \frac{1}{(2\pi)^{D/2}} \exp(-i\langle n, s \rangle)$ pour $n = (n_1, \dots, n_D)$.
- Sur $S = \mathbb{R}^D$, on peut considérer les bases d'ondelettes orthogonales (e.g. les ondelettes de Haar) dont on reparlera plus tard.

L'espace des mesures de Radon Soit $C_c(S)$ l'espace des fonctions continues qui s'annulent sur le bord de S (ou qui tendent vers 0 à l'infini si le domaine n'est pas borné). L'espace des mesures de Radon $\mathcal{X} = \mathcal{M}(S)$ est défini par

$$\mathcal{M}(S) \stackrel{\text{def}}{=} \{\text{formes linéaires continues sur } C_c(S)\}.$$

On rappelle que l'espace dual d'un espace vectoriel \mathcal{A} est défini comme l'ensemble des formes linéaires sur \mathcal{A} . Il est souvent noté \mathcal{A}^* . On peut munir l'espace $\mathcal{M}(S)$ d'une norme appelée (injustement) la variation totale, définie par

$$\|x\|_{\mathcal{M}(S)} = \sup_{f \in C_c(S), \|f\|_{\infty} \leq 1} \langle f, x \rangle.$$

On rappelle le résultat suivant

Proposition 1.1. Si $x \in L^1(S)$, alors on a $\|x\|_{\mathcal{M}(S)} = \|x\|_{L^1(S)} = \int_{s \in S} |x(s)| ds$.

Ainsi, la "variation totale" est aussi la "masse" de x . On préférera ce terme pour éviter les confusions avec une autre variation totale introduite ci-après.

L'intérêt principal de l'espace des mesures de Radon pour ce cours (et - de mon point de vue partial - en général) est qu'il est en un sens le plus petit espace fonctionnel qui inclut les masses de Dirac δ_s . En effet, pour donner un sens à l'évaluation ponctuelle $f(s) = \langle \delta_s, f \rangle$ pour tout s , il faut que f soit définie en tout point (ce n'est pas le cas des fonctions L^2 par exemple). Si on souhaite de plus que la forme linéaire δ_s soit continue (je ne rentre pas dans les détails topologiques ici), on doit ajouter la continuité des fonctions.

L'espace des fonctions à variation bornée Finalement, un espace qui a eu ses lettres de noblesse au début du vingtième siècle pour l'imagerie, mais qui perd progressivement en popularité est l'espace des fonctions à variation bornée $BV(S)$. Celui-ci est défini comme suit :

Définition 1.4 (Fonctions à variation bornée). L'espace $BV(S)$ des fonctions à variation bornée est l'ensemble des fonctions $x \in L^1(S)$ telles que la quantité suivante est bornée

$$TV(x) = \sup_{\phi \in (C_c^1(S))^D, \|\phi\|_{L^\infty(S)} \leq 1} \int_S x(s) \operatorname{div}(\phi)(s) ds < +\infty.$$

Cette définition peut paraître barbare et je pense que c'est le cas. D'un point de vue formel, on peut penser à cet espace comme l'espace des fonctions qui admettent des discontinuités le long de courbes suffisamment régulières. Une bonne façon d'imaginer les fonctions $x \in BV(S)$ est de considérer que leur "gradient" Dx est une mesure de Radon vectorielle dans $\mathcal{M}(S, \mathbb{R}^D)$.

Je reparlerai rapidement de cet espace plus tard dans le cours. Pour des gens intéressés par une introduction plus conséquente et ses applications en imagerie, je peux recommander l'excellent livre de G. Aubert et P. Kornprobst [1].

L'espace de Schwartz et les distributions tempérées On travaille ici sur l'espace $S = \mathbb{R}^D$.

Définition 1.5 (Espace de Schwartz). L'espace de Schwartz $\mathcal{S}(\mathbb{R}^D)$ est l'ensemble des fonctions à décroissance rapide, i.e. les fonctions $C^\infty(\mathbb{R}^D)$ dont toutes les dérivées sont à décroissance rapide, i.e. qui sont bornées quand on les multiplie par un polynôme d'ordre arbitraire.

Pour deux multi-indices α et β , on peut définir une norme $\|\cdot\|_{\alpha,\beta}$ par

$$\|f\|_{\alpha,\beta} = \|x^\alpha D^\beta f\|_\infty \quad (1.1)$$

où D^β est la dérivée d'ordre β . Ainsi $\mathcal{S}(\mathbb{R}^D) = \{f \in C^\infty(\mathbb{R}^D), \forall(\alpha, \beta), \|f\|_{\alpha,\beta} < +\infty\}$.

L'espace de Schwartz permet de définir l'ensemble des distributions tempérées.

Définition 1.6 (Distributions tempérées). L'ensemble des distributions tempérées est noté \mathcal{S}' . Il est défini comme l'ensemble des formes linéaires continues sur \mathcal{S} .

L'espace des distributions tempérées est gigantesque. En particulier, il contient $\mathcal{M}(\mathbb{R}^D)$. Ainsi, il contient par exemple les polynômes, les fonctions continues, les masses Dirac (tout comme $\mathcal{M}(\mathbb{R}^D)$). Il contient aussi d'autres éléments, comme les dérivées d'ordre arbitraire des Dirac (i.e. les opérateurs d'évaluation ponctuelle de toutes les dérivées partielles d'une fonction).

2 Quelques exemples de problèmes inverses linéaires

2.1 La déconvolution

Un prototype de problème inverse linéaire en imagerie est celui de la déconvolution : on dispose d'une image floue et on souhaite la rendre nette. Pour illustrer ce problème, nous allons d'abord nous placer dans un cadre continu. Je souhaite illustrer ici le fait que l'analyse mathématique et le travail dans des espaces de mesures permet de tirer des conclusions enrichissantes.

2.1.1 Une expérience numérique

On pose ainsi $S = \mathbb{R}^2$. Pour des raisons techniques qui deviendront claires très rapidement, on suppose que $x_0 \in C_0(S)$ est une image nette (en niveau de gris). En pratique, on utilisera l'image discrète de la figure 1, qu'on pourrait voir comme une fonction continue par interpolation. On suppose aussi qu'on dispose d'un noyau de convolution $h \in \mathcal{M}(S)$ ¹.

On suppose qu'on observe une image y définie par

$$y = h \star x_0 + b \quad (2.1)$$

où b est une perturbation et où \star représente le produit de convolution². Dans l'expérience numérique suivante, on considère un paramètre $\sigma > 0$ et on considère 3 choix distincts pour h .

— Un filtre de réplique :

$$h_0 = \frac{1}{3} (\delta_{0,\sigma} + \delta_{-\sigma/2,0} + \delta_{\sigma/2,0}) \quad (2.2)$$

— Un filtre de mouvement :

$$h_1 = \frac{1}{2\sigma} \delta_{([- \sigma, \sigma], 0)} \quad (2.3)$$

où $\delta_{([- \sigma, \sigma], 0)}$ est la mesure uniforme sur le segment $s = \{(s_1, 0), s_1 \in [-\sigma, \sigma]\}$.

— Un filtre gaussien :

$$h_2(s) = \frac{1}{2\pi\sigma} \exp\left(-\frac{\|s\|_2^2}{2\sigma^2}\right). \quad (2.4)$$

La figure 2 montre ces trois filtres, le résultat de la convolution pour ces 3 filtres, ainsi que le résultat d'une déconvolution par une méthode numérique standard. Comme on peut le voir ici, les valeurs de SNR (Signal-to-Noise Ratio) des différents résultats de reconstruction sont très distincts. Ce qui est surprenant, c'est qu'à première vue, les images qui semblent le plus difficile à reconstruire (au moins à partir de leur SNR) sont les mieux reconstruites. Comment expliquer ce phénomène ?

1. Une image n'appartient que très rarement à $C_0(S)$, car elle possède des discontinuités. Au contraire, vous verrez dans d'autres cours qu'un noyau de convolution réaliste est en général C^∞ . Par la suite on intervertira les deux espaces, mais j'ai besoin de ces hypothèses pour cet exemple introductif.

2. Définir proprement la convolution pour des mesures est un petit peu délicat. Pour les lecteurs intéressés par un approfondissement, je recommande chaudement le livre de Y. Katznelson [16]



FIGURE 1 – L'image de référence utilisée pour ce cours.

2.1.2 Analyse des résultats

Pour analyser ces résultats, on va utiliser la fameuse formule de diagonalisation des opérateurs de convolution par la transformation de Fourier \mathcal{F} . Dans le domaine fréquentiel, l'équation (2.1) peut être réécrite sous la forme :

$$\mathcal{F}y = \mathcal{F}h \cdot \mathcal{F}x_0 + \mathcal{F}b. \quad (2.5)$$

où on définit $\mathcal{F}x : \mathbb{R}^D \rightarrow \mathbb{C}$ par

$$(\mathcal{F}x)(\xi) = \int_{\mathbb{R}^D} x(s) \exp(-i\langle s, \xi \rangle) dt. \quad (2.6)$$

Pour effectuer une déconvolution, on peut simplement utiliser diviser $\mathcal{F}y$ par $\mathcal{F}h$, ce qui pose des soucis numériques lorsque $\mathcal{F}h$ s'annule. Dans l'expérience précédente, nous avons donc préféré résoudre un problème discrétisé et régularisé de la forme suivante :

$$\min_{x \in \mathbb{R}^N} \frac{1}{2} \|y - h \star x\|_2^2 + \frac{\lambda}{2} \|x\|_2^2 \quad (2.7)$$

où $\lambda > 0$ est un paramètre de régularisation. La solution \hat{x} de ce problème est donnée par

$$\hat{x} = \mathcal{F}^{-1} \left(\frac{\bar{\mathcal{F}}h \cdot \mathcal{F}y}{|\mathcal{F}h|^2 + \lambda} \right).$$

Dans le cas où $\mathcal{F}h(\xi) \neq 0$ pour tout ξ , on peut prendre $\lambda = 0$ et on obtient :

$$\hat{x} = x_0 + \mathcal{F}^{-1} \left(\mathcal{F}b \cdot \frac{\bar{\mathcal{F}}h}{|\mathcal{F}h|^2} \right). \quad (2.8)$$

On voit ici le rôle crucial que joue la décroissance de $|\mathcal{F}h|$: plus il décroît rapidement, plus les hautes fréquences du bruit b vont être amplifiées. Peut-on capturer ce phénomène de décroissance mathématiquement? La réponse est positive et fait appel à des ouvrages d'analyse harmoniques utilisés (de façon surprenante) surtout en théorie de nombres [15].

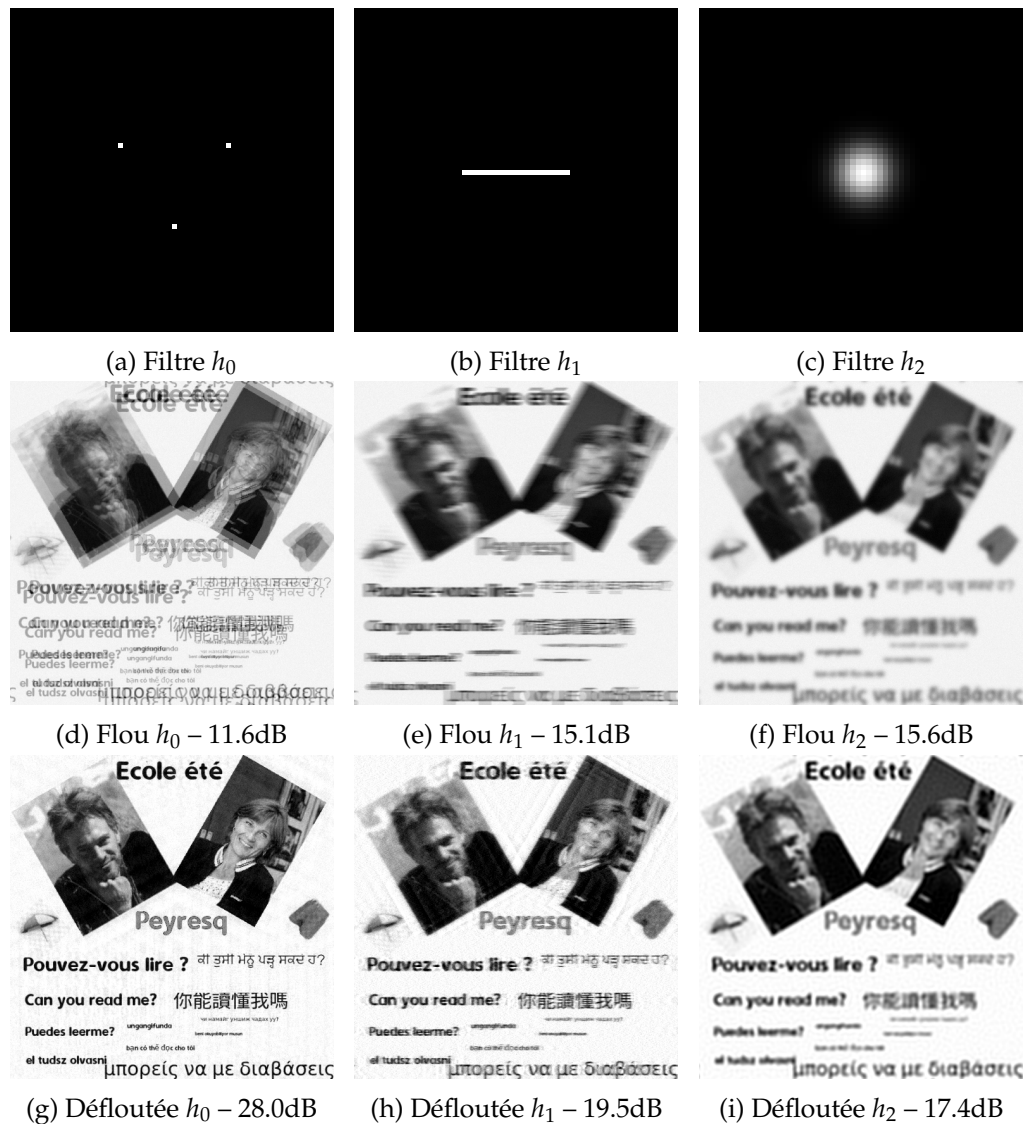


FIGURE 2 – Une expérience introductive : convolution et déconvolution pour 3 filtres différents. Comme on le voit, le qualité de la déconvolution dépend fortement du filtre utilisé. Le SNR de l'image à traiter ne semble pas donner d'indication pertinente sur la qualité du traitement.

2.1.3 Régularité et décroissance de la transformée de Fourier

Sans rentrer dans trop de détails techniques qu'un lecteur intéressé pourra explorer par lui-même, je souhaite donner deux résultats qui me semblent importants. Le premier résultat est énoncé sur le tore $S = \mathbb{T} = \mathbb{R} \setminus 2\pi\mathbb{Z}$. Dans ce cas, on peut définir aisément une série de Fourier pour tout $x \in L^1(S)$ [16]. Le n -ième terme de la série de Fourier est défini par $(\mathcal{F}x)[n] = \int \exp(-i \cdot ns)x(s) ds$.

Théorème 2.1 (Décroissance sur les classes régulières [16]).

— Si h est k fois différentiable et que $h^{(k-1)}$ est absolument continue, alors

$$|(\mathcal{F}h)[n]| \leq \min_{0 \leq j \leq k} \frac{\|h^{(j)}\|_{L^1}}{|n|^j}. \quad (2.9)$$

— Si $h \in BV(S)$, alors

$$|(\mathcal{F}h)[n]| \leq \frac{TV(h)}{|n|}. \quad (2.10)$$

Ce théorème indique que plus un noyau est régulier, plus sa transformée de Fourier décroît. De plus, les estimées ci-dessus sont ajustées, au sens où il existe des fonctions qui atteignent les bornes supérieures. Que peut-on dire dans le cas où h est une mesure singulière? Premièrement, le cas du filtre h_0 défini comme une somme de masses de Dirac est facile à traiter : sa transformée de Fourier est une somme finie d'exponentielles complexes, qui ne décroît pas à l'infini. On peut le voir sur la figure 3a.

Qu'en est-il des mesures supportées sur des courbes? Cette question est particulièrement importante en imagerie, puisque les mesures supportées sur des courbes décrivent les flous de mouvement.

Théorème 2.2 (Décroissance sur les mesures [15]). Soit $\gamma : [0, 1] \rightarrow \mathbb{R}^2$ une courbe continue. Soit μ_γ la mesure définie pour toute fonction f continue par

$$\langle \mu_\gamma, f \rangle = \int_{s \in [0, 1]} f(\gamma(s)) ds. \quad (2.11)$$

Si les vecteurs dérivées $\gamma'(s)$ et $\gamma''(s)$ sont linéairement indépendants pour tout s . Alors

$$\max_{\|\xi\|_2=n} |(\mathcal{F}\mu_\gamma)(\xi)| = O(n^{-1/2}). \quad (2.12)$$

Le taux de décroissance indiqué ci-dessus est bien plus lent que ceux établis pour des fonctions régulières. De plus, des inégalités opposées du type $\int_{\xi, \|\xi\|_2=n} |(\mathcal{F}\mu_\gamma)(\xi)| d\xi \geq cn^{-\alpha}$ pour un certain $\alpha > 0$ et $c > 0$ sont aussi disponibles (voir [15, P194]) et indiquent que la transformée de Fourier d'une mesure supportée sur une courbe *ne peut pas décroître trop rapidement*. Le théorème précédent est une conséquence assez directe du lemme de Van der Corput sur les intégrales oscillantes.

Pour revenir à notre exemple, ces résultats sont capturés dans la figure 3 : on voit que plus les noyaux sont singuliers, moins leur transformée de Fourier contient de valeurs proches de 0. En particulier, le premier noyau ne décroît absolument pas à l'infini. Le second est invariant dans la direction verticale et il décroît lentement dans la direction orthogonale. Le troisième est terrible : la transformée d'une gaussienne est une gaussienne, et les hautes fréquences sont irrémédiablement perdues.

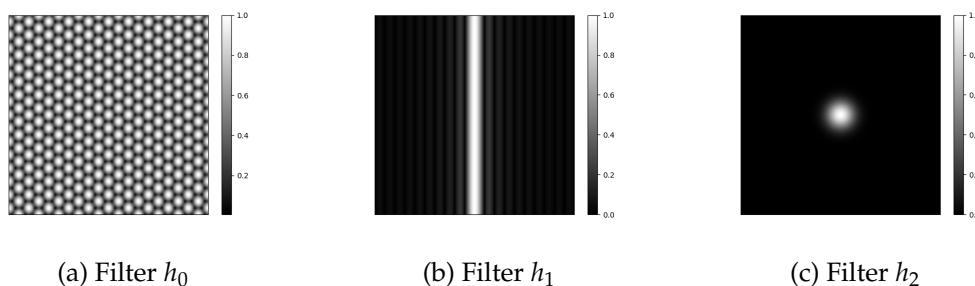


FIGURE 3 – Transformée de Fourier des 3 différents filtres. La vitesse de décroissance du noyau dépend fortement.

2.1.4 Conclusions

A travers cette expérience, plusieurs points peuvent être retenus :

- La régularité du filtre de convolution joue un rôle crucial sur sa décroissance dans le domaine de Fourier et sur la difficulté d'un problème de déconvolution.
- Ainsi, les filtres de mouvement (vision par ordinateur) sont bien plus faciles à inverser que les filtres passe-bas de l'optique conventionnelle (e.g. microscopie).
- Le fait de travailler sur des espaces de mesures a permis d'utiliser des outils d'analyse harmonique assez féconds pour comprendre le phénomène.

2.2 Défloutage vs déconvolution

Il est fréquent de voir apparaître le terme déconvolution pour un problème où on souhaite rendre une image nette. De mon point de vue, ce terme devrait être utilisé avec plus de précautions. En effet, un opérateur de convolution possède une PSF invariante spatialement, alors que la majorité des systèmes optiques possèdent des réponses impulsionnelles qui varient dans le champs de vue.

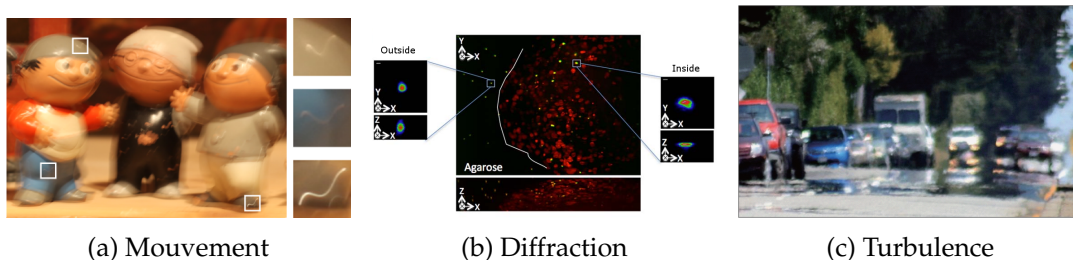


FIGURE 4 – Trois exemples d'images floues. Observer comme dans chaque cas, on observe que le flou est non stationnaire.

Les figures 4 et 5 montrent de façon indirecte trois opérateurs de "floutage" différents. En particulier, la représentation 5 montre l'application d'un opérateur à

une grille de masses de Dirac. On peut voir que les réponses impulsionnelles varient dans le champs de vue. Si on souhaite rendre ces images plus nettes, on ne peut pas parler de déconvolution et il faut³ employer un terme différent tel que défloutage. En particulier le terme “space varying deconvolution” devrait être proscrit puisqu’il contient deux termes antagonistes. On peut par contre utiliser le terme “space varying deblurring”.

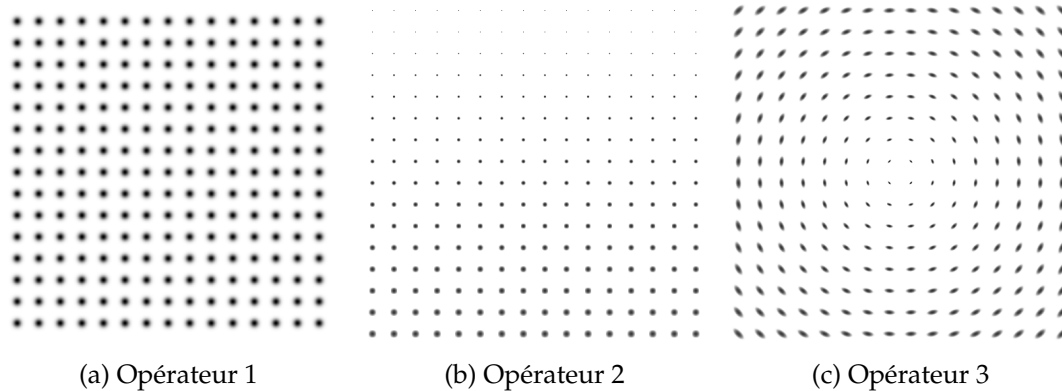


FIGURE 5 – Trois exemples d’opérateurs régularisants. On les affiche ici en les appliquant à une grille de masses de Dirac. Le premier est stationnaire : la PSF est invariante spatialement. Le second et le troisième sont instationnaires, ce ne sont pas des opérateurs de convolution !

2.3 Super-résolution en microscopie optique/SMLM

Dans l’exemple précédent, nous avons vu qu’introduire le formalisme de la théorie de la mesure permettait de prédire (ou au moins de formaliser) la difficulté de tâches de déconvolution. Dans certains cas, c’est aussi le cadre naturel dans lequel les signaux vivent. C’est le cas de la super-résolution par localisation de molécules uniques, qui est un domaine important de la microscopie optique (prix Nobel 2014).

Dans cette modalité d’imagerie, uniquement quelques molécules sont activées pour produire une image. Ceci permet de déterminer leur localisation avec une précision sous-pixellique avec des algorithmes adaptés. On peut utiliser des milliers d’images de ce type (des molécules différentes sont activées à chaque fois) et accumuler les positions retrouvées. Ceci permet de multiplier les résolutions par des facteurs de l’ordre de 10. Le principe est illustré sur la figure 6.

Comment modéliser ceci ? Si on néglige la taille des molécules fluorescentes⁴, on peut poser

$$x_0 = \sum_{p=1}^P \alpha_p \delta_{s_p},$$

3. de mon point de vue et malgré une terminologie différente dans de nombreux articles

4. La majorité des fluorophores (les molécules fluorescentes utilisées pour la localisation) sont constituées de 20 à 100 atomes, ce qui résulte en une taille de molécules variant entre 2 et 10 nanomètres. Les résolutions d’image maximales atteignables aujourd’hui sont justement de cet ordre de grandeur, et on peut donc commencer à questionner le modèle du Dirac... Si vous avez des informations à partager, n’hésitez pas !

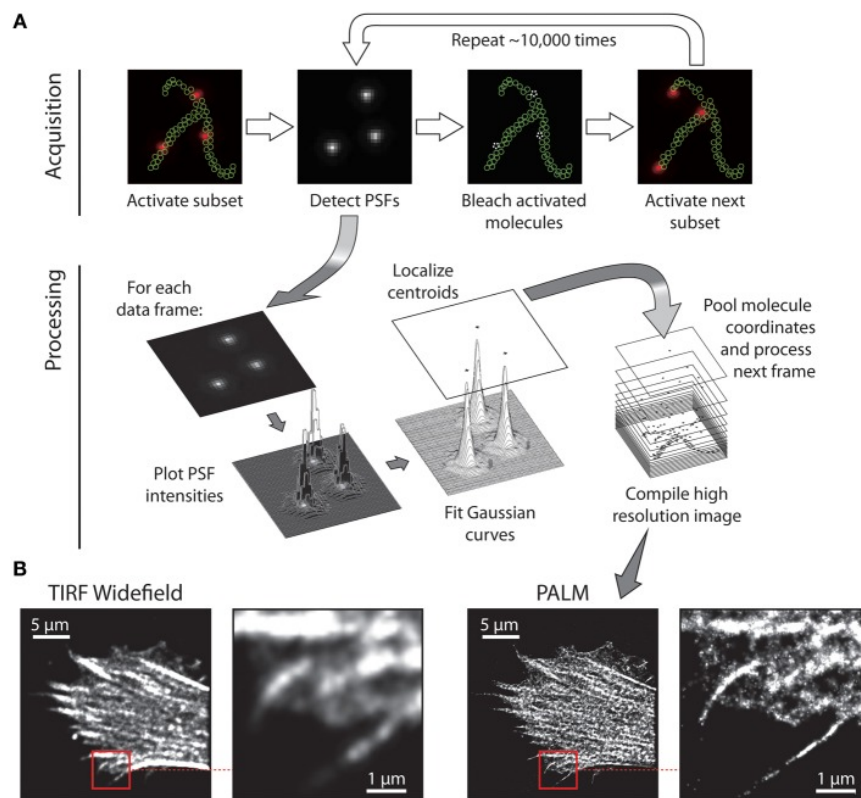


FIGURE 6 – Le principe de la microscopie par localisation de molécules uniques (SMLM).

où $s_p \in \mathbb{R}^D$ est la position d'une molécule et $\alpha_p > 0$ est son intensité de fluorescence. Ainsi, une image $x_0 \in \mathcal{M}(\mathbb{R}^D)$ est modélisée comme une mesure de Radon.

Un modèle simple d'acquisition en imagerie est le suivant, on observe M intensités lumineuses y_m en issues de points $z_m \in \mathbb{R}^D$. Ainsi :

$$\begin{aligned} y_m &= \mathcal{P}((h \star x)(z_m)) \\ &= \mathcal{P}(\langle x, h(\cdot - z_m) \rangle), \end{aligned}$$

où \mathcal{P} est une perturbation aléatoire ou déterministe telle que l'ajout de bruit gaussien et la quantification. Le problème inverse qu'on se pose ici est le suivant :

Connaissant $y \in \mathbb{R}^M$, retrouver les positions s_p et intensités α_p .

Nous verrons plus tard quelques solutions efficaces pour ce problème inverse reposant sur l'optimisation sur des espaces de mesures (dans ce cas, la programmation semi-infinie). Là-encore, le formalisme de la théorie de la mesure est fécond.

2.4 L'échantillonnage compressé

Dans de nombreux imageurs, il est tentant de réduire le nombre de mesures effectuées, pour plusieurs raisons telles que :

- Réduire les temps d'acquisition,
- Réduire les artefacts de mouvements (conséquence du premier effet),
- Réduire les effets délétères tels que le photobleaching en microscopie, la toxicité en tomographie par rayons X ou l'inconfort sur des imageurs par résonance magnétique,
- Réduire les coûts de fabrication (e.g. en réduisant le nombre de capteurs CCD).

Un exemple qui illustre assez parfaitement ces motivations est l'imagerie par résonance magnétique (IRM). Dans cette modalité, on peut mesurer les valeurs de la transformée de Fourier d'une image $x \in \mathcal{M}(S)$, où S est un sous-domaine de \mathbb{R}^2 ou \mathbb{R}^3 . On dispose ainsi d'un ensemble de mesures $y = (y_m)_{1 \leq m \leq M}$ telles que

$$y_m = \mathcal{P}((\mathcal{F}x)(z_m)),$$

où $(z_m)_{1 \leq m \leq M}$ est un ensemble de positions dans \mathbb{R}^D qu'on peut choisir assez librement⁵. Pour réduire les temps d'acquisition, on peut se poser le problème suivant :

Pour une erreur de reconstruction fixée, trouver le nombre de mesures M minimum, l'ensemble des positions z_m et un algorithme de reconstruction qui permettent de retrouver un signal x_0 , ou plus généralement d'une classe de signaux \mathcal{X} ?

Cette question est le coeur de ce qu'on appelle aujourd'hui l'échantillonnage compressé. Nous donnerons quelques éléments de compréhension de ce domaine, plus tard dans ce cours, .

2.5 Le cadre de travail général

J'espère avoir montré à travers ces quelques exemples, qu'il est intéressant de travailler dans un cadre abstrait, pour pouvoir traiter les problèmes inverses de manière unifiée.

5. Le modèle complet est plus complexe, mais je ne souhaite pas rentrer dans les détails ici.

2.5.1 Les problèmes inverses linéaires

Soit $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ et $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ deux espaces vectoriels normés de dimension finie ou infinie. Si les espaces sont de dimension finie, on posera dans tout le cours $\dim(\mathcal{X}) = N$ et $\dim(\mathcal{Y}) = M$. L'objet d'étude central dans un problème inverse linéaire est un opérateur linéaire $A : \mathcal{X} \rightarrow \mathcal{Y}$. On notera $\mathcal{L}(\mathcal{X}, \mathcal{Y})$ l'ensemble des opérateurs linéaires de \mathcal{X} dans \mathcal{Y} .

Définition 2.1 (Problème inverse linéaire). Un problème inverse linéaire consiste à estimer un vecteur $x_0 \in \mathcal{X}$ à partir d'un opérateur linéaire $A : \mathcal{X} \rightarrow \mathcal{Y}$ connu et de mesures $y \in \mathcal{Y}$ de la forme :

$$y = Ax_0. \quad (2.13)$$

Les mesures effectuées par des appareils numériques subissent des perturbations déterministes (e.g. quantification par des convertisseurs ADC) ou stochastiques (e.g. arrivée aléatoire de photons sur les capteurs). Il est essentiel de prendre en compte cet effet et d'étudier la stabilité des méthodes de résolution à des perturbations. Ainsi, on considèrera plutôt des modèles de dégradation du type :

$$y = \mathcal{P}(Ax_0) \quad (2.14)$$

où $\mathcal{P} : \mathcal{Y} \rightarrow \mathcal{Y}$ modélise des détériorations du signal mesuré.

Si on doit retrouver x à partir de y sans connaître l'opérateur A on parle de problème inverse aveugle.

Définition 2.2 (Problème inverse aveugle). Un problème inverse aveugle consiste à retrouver $x_0 \in \mathcal{X}$ à partir de $y = Ax_0 \in \mathcal{Y}$ sans connaître ni le signal $x_0 \in \mathcal{X}$ ni l'opérateur A .

Les problèmes inverses aveugles sont un domaine de recherche en grande partie ouvert et nous ne l'évoquerons pas ou seulement très partiellement dans ce cours.

2.5.2 Difficultés

Trois difficultés principales se posent lorsqu'on étudie un problème inverse linéaire. Elles ont été bien identifiées par Jacques Hadamard⁶.

i) Existence Un système linéaire n'admet pas toujours de solution. L'ajout de bruit notamment peut faire que $y \notin \text{ran}(A)$. Ce problème n'est pas en général le plus délicat à traiter. En effet, une solution simple et souvent efficace consiste simplement à chercher une solution qui minimise le critère suivant :

$$\inf_{x \in \mathcal{X}} \frac{1}{2} \|Ax - y\|_{\mathcal{Y}}^2, \quad (2.15)$$

où $\|\cdot\|_{\mathcal{Y}}$ est une norme sur l'espace \mathcal{Y} . Lorsqu'on travaille en dimension finie, on peut montrer que le problème ci-dessus admet toujours une solution et on a donc résolu le problème d'existence. Des précautions/conditions supplémentaires sont nécessaires en dimension infinie, mais sauf cas pathologiques, tout se passe bien. Nous ne rentrerons pas dans ces détails dans ce cours.

6. 1865-1963, médaille d'or du CNRS, président de la SMF, père de la femme de Dreyfus et bien d'autres encore...

ii) Unicité Un souci souvent bien plus délicat à traiter est celui de l'unicité. Etant donné $y \in \mathcal{Y}$, il peut exister plusieurs $x \in \mathcal{X}$ tels que $Ax = y$. C'est le cas dès que le noyau de A n'est pas restreint au singleton 0 ($\ker(A) \neq \{0\}$). Comment en sélectionner un ? On peut notamment penser ici à l'échantillonnage compressé où toute la difficulté est de sélectionner une solution parmi un ensemble énorme.

Une solution fréquente est de sélectionner la pseudo-inverse qu'on peut définir (dans des espaces de Hilbert) comme la limite suivante :

$$\lim_{\tau \rightarrow 0} \operatorname{argmin}_{x \in \mathcal{X}} \frac{1}{2} \|Ax - y\|_{\mathcal{Y}}^2 + \frac{\tau}{2} \|x\|_{\mathcal{X}}^2. \quad (2.16)$$

Sous des hypothèses raisonnables (e.g. dimension finie), l'ajout du terme $\|x\|_{\mathcal{X}}^2$ assure que la fonctionnelle ci-dessus est fortement convexe et garantit l'unicité du minimiseur. Ce choix est fréquemment discutable et on verra des techniques de sélection de solution plus avancées et efficaces par la suite.

iii) Stabilité Finalement, un problème fréquent et difficile est celui de la stabilité au bruit. Etant donné des mesures du type :

$$y = Ax_0 + b, \quad (2.17)$$

où $b \in \mathcal{Y}$ est une perturbation, on souhaite qu'un algorithme nous renvoie une estimée \hat{x} proche de x_0 , au sens où :

$$\|\hat{x} - x_0\|_{\mathcal{X}} \leq C \|b\|_{\mathcal{Y}}, \quad (2.18)$$

pour une certaine constante $C \geq 0$. Cette condition de stabilité garantira en particulier que $\hat{x} = x_0$ dans le cas sans bruit correspondant à $\|b\|_{\mathcal{Y}} = 0$.

A retenir

Dans ce cours, nous travaillerons essentiellement sur des problèmes inverses mal posés, c'est-à-dire des problèmes pour lesquels un des trois points ci-dessus ne sont pas vérifiés. En sens, les systèmes d'acquisition numériques, mènent toujours à des problèmes inverses mal posés puisqu'on observe un signal continu avec un nombre fini de mesures. En imagerie computationnelle, il est important de concevoir un système pour que la reconstruction soit un problème le mieux posé possible.

3 Opérateurs intégraux

Aujourd'hui, la très grande majorité des systèmes d'acquisition sont numériques. Dans ce cas, on peut identifier \mathcal{Y} à \mathbb{R}^M ou \mathbb{C}^M , où M représente un nombre de mesures. N'importe quel opérateur linéaire A de \mathcal{X} dans \mathbb{K}^M peut s'écrire sous la forme :

$$y_m = (Ax)[m] = \langle a_m, x \rangle, \quad (3.1)$$

où $a_m \in \mathcal{X}^*$ est une forme linéaire sur \mathcal{X} . Pour les gens habitués à travailler en dimension finie, on peut donc voir l'opérateur A comme une "matrice" (potentiellement infinie suivant une direction) du type⁷ :

$$A = \begin{pmatrix} a_1 \\ \vdots \\ a_M \end{pmatrix}.$$

Donnons tout de suite 2 exemples :

Super-résolution dans ce cas, on a vu précédemment que $a_m = h(\cdot - z_m)$, où z_m est une position d'échantillonnage et h un noyau de convolution.

IRM on peut ici poser $a_m = \exp(-2i\langle \cdot, z_m \rangle)$ où z_m est une fréquence à mesurer.

Remarque 3.1. On peut ici faire un lien entre les problèmes inverses linéaire et un problème important d'analyse appelé le *problème des moments*. En effet, des mesures du type (3.1) sont appelés *moments généralisés* de x . Si ψ_m est un monôme, on évalue un moment usuel (e.g. moyenne, centre de masse, écart-type). Si ψ_m est une exponentielle complexe (cas de l'IRM), on mesure un moment trigonométrique. De nombreuses techniques numériques existent pour le problème des moments et il y a donc des liens à tisser entre les deux domaines.

3.1 Définitions

La vision d'un opérateur linéaire comme une collection de formes linéaires (a_1, \dots, a_M) est souvent suffisante. Cependant, en imagerie computationnelle, il est fréquemment fécond de voir un opérateur linéaire comme le produit d'un *opérateur linéaire intégral* et d'un *opérateur d'échantillonnage*. Nous introduisons ces deux objets ci-dessous.

Définition 3.1 (Opérateur d'échantillonnage). Soit $Z = (z_1, \dots, z_M)$ un ensemble un ensemble de M points de $S \subset \mathbb{R}^D$. L'opérateur d'échantillonnage $\mathbb{I}_Z : C(S) \rightarrow \mathbb{K}^M$ est défini par

$$\begin{array}{ccc} \mathbb{I}_Z : C(S) & & \rightarrow \mathbb{K}^M \\ f & & \mapsto (f(z_1), \dots, f(z_M)). \end{array}$$

7. De mon point de vue il serait probablement préférable d'écrire $y = A^*x$ avec $A^* : \mathcal{X} \rightarrow \mathbb{K}^M$ et $A = [a_1, \dots, a_M]$. Ce choix a plus de sens du point de vue de l'analyse fonctionnelle où on écrit souvent les formes linéaires avec le symbole $*$ (voir par exemple l'excellent livre de Ekeland et Temam [9]). Malheureusement pour moi, ce choix est très rare dans la littérature et j'ai donc choisi de me conformer à la masse pour ne pas perdre mes lecteurs.

On introduit aussi les opérateurs linéaires intégraux H de façon formelle ci-dessous.

Définition 3.2 (Opérateur linéaire intégral). Un opérateur linéaire intégral est une application linéaire d'un espace de fonctions mesurables \mathcal{X} sur un domaine S dans un espace de fonctions \mathcal{Y} défini sur un domaine T . Elle est définie par une expression du type :

$$(Hx)(t) = \int_{s \in S} k(t, s) dx(s) = \langle k(t, \cdot), x \rangle, \quad (3.2)$$

La fonction (généralisée) $k : T \times S \rightarrow \mathbb{K}$ est appelée *noyau*⁸ de l'opérateur H .

Dans tous les exemples introductifs et bien d'autres encore, on peut écrire l'opérateur A sous la forme du produit

$$A = \text{III}_Z \cdot H. \quad (3.3)$$

On peut se poser la question à quel point il est restrictif pour un opérateur linéaire de prendre la forme intégrale (3.2). Une réponse a été apportée par Laurent Schwartz en 1952. Si on autorise le noyau k à vivre dans l'espace de Schwartz, essentiellement tous les opérateurs linéaires (par exemple une dérivée ou un laplacien) peuvent s'écrire sous cette forme !

Théorème 3.1 (Théorème du noyau intégral de Schwartz [14]). Soient S et T deux ensembles ouverts de \mathbb{R}^D . Toute distribution tempérée $k \in \mathcal{S}'(T \times S)$ définit une application linéaire continue $H : \mathcal{S}(S) \rightarrow \mathcal{S}'(T)$ telle que

$$\langle k, u \otimes v \rangle = \langle Hv, u \rangle \quad \text{pour tout } u \in \mathcal{S}(S) \text{ et } v \in \mathcal{S}(T). \quad (3.4)$$

Réciproquement, pour toute application linéaire continue H , il existe une unique distribution $k \in \mathcal{S}'(T \times S)$ telle que (3.4) soit valide.

Ce théorème me semble fondamental, car il fait écho au résultat bien connu en dimension finie, qui assure que toute application linéaire peut être codée par une matrice. On peut ainsi rapprocher le noyau d'un opérateur intégral à la représentation matricielle d'une application linéaire.

Par contre, il faut bien noter que l'opérateur III_Z est défini uniquement sur les fonctions continues et il faut donc prendre des précautions pour que le produit $\text{III}_Z \cdot H$ ait un sens.

Donnons quelques exemples d'opérateur intégraux :

IRM, paragraphe 2.1 on échantillonne la transformée de Fourier dont le noyau est défini pour tout $s \in \mathbb{R}^D$ par

$$k(t, s) = \exp(-i\langle s, t \rangle) \quad (3.5)$$

Computerized Tomography on échantillonne la transformée de Radon définie pour tout $t = (t_1, t_2) \in [0, \pi] \times \mathbb{R}$ par

$$k(t, \cdot) = \delta_{L_{t_1, t_2}} \quad (3.6)$$

où $\delta_{L_{t_1, t_2}}$ est la mesure uniforme sur la ligne L_{t_1, t_2} d'angle t_1 et de décalage à l'origine de valeur t_2 .

8. D'où le symbole k qui se réfère à "kernel" en anglais.

Super-résolution, paragraphe 2.3 on échantillonne un produit de convolution dont le noyau intégral est défini par

$$k(t, s) = h(s - t), \quad (3.7)$$

pour un certain filtre de convolution h .

Définition 3.3 (Réponse impulsionnelle et fonction d'étalement du point⁹). La *réponse impulsionnelle* d'un opérateur H de noyau k continu en un point s est la fonction $H\delta_s = k(\cdot, s)$.

La *fonction d'étalement du point* (ou PSF pour Point Spread Function) en un point $s \in S$ est la fonction $k(\cdot + s, s)$, i.e. la réponse impulsionnelle recentrée en 0.

En particulier, dans le cas d'un opérateur de convolution par un filtre h , on a $k(t, s) = h(s - t)$ et la PSF en un point s est simplement $k(\cdot + s, s) = h(\cdot)$, i.e. elle ne varie pas en fonction de s . Notons bien qu'il est nécessaire que le noyau k soit continu pour pouvoir donner un sens au produit $H\delta_s$.

Proposition 3.1 (Produit d'opérateurs intégraux). Soient $H_1 \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ et $H_2 \in \mathcal{L}(\mathcal{Y}, \mathcal{Z})$ deux opérateurs intégraux de noyaux respectifs k_1 et k_2 dans $L^2(T \times S)$ et $L^2(U \times T)$ où U est un espace mesuré et \mathcal{Z} un espace de fonctions mesurables sur U .

Alors l'opérateur $H = H_2 \circ H_1$ est un opérateur intégral linéaire de $\mathcal{L}(\mathcal{X}, \mathcal{Z})$ et son noyau intégral k est donné pour tout (u, s) par

$$k(u, s) = \int_{t \in T} k_2(u, t) k_1(t, s) ds. \quad (3.8)$$

Démonstration. Exercice. □

Remarquez que ce résultat est exactement le pendant continu de la formule du produit matrice-matrice.

3.2 Effet régularisant

Une propriété importante des opérateurs intégraux est leur aspect régularisant. Par exemple vous avez dû voir un jour que (sous des hypothèses raisonnables) la fonction $g = f \star h$ possède (au moins) la régularité de la fonction la plus régulière entre f et g . Pour un opérateur intégral plus général, on peut par exemple établir le résultat suivant.

Proposition 3.2 (Effet régularisant d'un opérateur intégral). Soit H un opérateur intégral de noyau $k \in W^{R, \infty}(T \times S)$ avec $R \in \mathbb{N}$, c'est-à-dire que toutes les dérivées partielles de k jusqu'à l'ordre R sont bornées. Alors pour $\mu \in \mathcal{M}(S)$, on a

$$H\mu \in W^{R, \infty}(T). \quad (3.9)$$

Démonstration. On se place dans le cas où S et T sont des intervalles (1D) par simplicité. La preuve dans le cas multi-dimensionnelle est identique.

Pour tout t , on a $H\mu(t) = \int_{s \in S} k(t, s) d\mu(s) \leq \|k(t, \cdot)\|_{L^\infty(S)} \|\mu\|_{\mathcal{M}(S)}$. Donc $H\mu$ est une fonction bornée dans $L^\infty(T)$ et on a montré le théorème pour $R = 0$.

9. La majorité des sources que j'ai pu consulter (e.g. Wikipédia) restent assez évasives sur ce concept, sans définition mathématique précise.

Soit (δ_i) une suite tendant vers 0. Pour tout $t \in T$ et $i \in \mathbb{N}$, on a

$$\frac{H\mu(t + \delta_i) - H\mu(t)}{\delta_i} = \int_{s \in S} \frac{k(t + \delta_i, s) - k(t, s)}{\delta_i} d\mu(s). \quad (3.10)$$

Comme $k \in W^{R, \infty}(S \times T)$, elle est en particulier Lipschitz et $|k(t + \delta_i, s) - k(t, s)| \leq L\delta_i$ pour un certain $L > 0$. Ainsi les fractions $\frac{k(t + \delta_i, s) - k(t, s)}{\delta_i}$ sont uniformément bornées et on peut appliquer le théorème de convergence dominée de Lebesgue :

$$\begin{aligned} \lim_{i \rightarrow \infty} \frac{H\mu(t + \delta_i) - H\mu(t)}{\delta_i} &= \int_{s \in S} \lim_{i \rightarrow \infty} \frac{k(t + \delta_i, s) - k(t, s)}{\delta_i} d\mu(s) \\ &= \int_{s \in S} \partial_1 k(t, s) d\mu(s) \leq \|\partial_1 k(t, s)\|_{L^\infty(S)} \|\mu\|_{\mathcal{M}(S)}. \end{aligned}$$

Le théorème est montré pour $R = 1$. Par récurrence, on obtient le théorème pour tout R . \square

Ce résultat est à mettre en regard des théorèmes 2.1 sur la décroissance en Fourier des fonctions de classe C^R . On voit ainsi que plus le noyau intégral k est lisse, plus il va atténuer les hautes fréquences de la mesure x et plus il sera difficile de les retrouver!

Remarque 3.2. Le théorème ci-dessus est donné dans un cadre très restreint et peut être très largement affiné. Par exemple, dans la preuve, on a utilisé uniquement le fait que k soit dérivable par rapport à la première variable. On peut aussi faire l'hypothèse que k est régulière suivant la deuxième variable, ce qui donnerait une décomposition tensorielle différente. De même, l'espace $W^{R, \infty}(S \times T)$ est régulier dans le produit des variables (s, t) . En réalité, on pourrait considérer des espaces fonctionnels "mixtes" où k a une certaine régularité suivant la variable s à t fixé, et une autre régularité suivant la variable t à s fixé. Nous ne rentrons pas dans ces détails ici pour .

A retenir

Plus le noyau intégral d'un opérateur H est lisse, plus l'opérateur est régularisant, et plus il sera difficile de recomposer les hautes fréquences (i.e. la résolution) du signal x sous-jacent à partir des mesures $y = \mathcal{P}(\mathbb{I}_Z Hx)$! Si possible, il faut donc essayer de construire des systèmes de mesure avec des noyaux singuliers (e.g. computerized tomography, flous de mouvement).

3.3 Opérateurs de Hilbert-Schmidt

Nous ne nous sommes pas embarrassés ici de conditions techniques sur les espaces fonctionnels \mathcal{X} et \mathcal{Y} et sur le noyau k . Cette question est loin d'être évidente. Une solution relativement simple que nous adopterons pour commencer est de travailler sur des espaces de Hilbert. Nous ferons ici le choix de poser $\mathcal{X} = L^2(T)$ et $\mathcal{Y} = L^2(S)$ où S et T sont typiquement des sous-ensembles de \mathbb{R}^D ¹⁰ ou \mathbb{T}^D . On notera (e_n) une base hilbertienne de \mathcal{X} et (f_m) une base hilbertienne de \mathcal{Y} . Ainsi tout

10. La dimension D peut ne pas être identique pour T et S , même si nous ne rencontrerons pas de tels opérateurs dans ce cours

$x \in \mathcal{X}$ peut être décomposé sous la forme $x = \sum_{n=1}^{\infty} \langle x, e_n \rangle e_n$ et tout $y \in \mathcal{Y}$ sous la forme $y = \sum_{m=1}^{\infty} \langle y, f_m \rangle f_m$. Les signaux x et y sont entièrement caractérisés par les séries associées (α_n) et (β_n) .

Définition 3.4 (Norme de Hilbert-Schmidt). La norme de Hilbert-Schmidt d'un opérateur $H : L^2(S) \rightarrow L^2(T)$ est définie par

$$\|k\|_{HS}^2 \stackrel{\text{def}}{=} \|k\|_{L^2(T \times S)}^2 = \int_{s \in S} \int_{t \in T} k^2(t, s) ds dt. \quad (3.11)$$

Il ne faut pas confondre cette norme avec une autre norme très populaire :

Définition 3.5 (Norme spectrale). La norme spectrale d'un opérateur intégral $H : L^2(S) \rightarrow L^2(T)$ est définie par

$$\|H\|_{2 \rightarrow 2}^2 \stackrel{\text{def}}{=} \sup_{\|x\|_{L^2(S)} \leq 1} \|Hx\|_{L^2(T)}^2 \quad (3.12)$$

Définition 3.6 (Opérateur de Hilbert-Schmidt). Un opérateur de Hilbert-Schmidt H est un opérateur intégral $H : L^2(S) \rightarrow L^2(S)$ ¹¹ dont la norme de Hilbert-Schmidt est bornée $\|H\|_{HS}^2 < +\infty$.

Nous traiterons ci-dessous le cas plus général d'un opérateur intégral ayant un domaine d'entrée et de sortie différents $H : \mathcal{X} \rightarrow \mathcal{Y}$ et une norme de Hilbert-Schmidt bornée :

$$\|H\|_{HS} < +\infty. \quad (3.13)$$

Un tel opérateur n'est donc pas à proprement parler un opérateur de Hilbert-Schmidt. La dernière condition implique en particulier que la *norme spectrale* de H est bornée. En effet, en utilisant l'inégalité de Cauchy-Schwarz, on obtient

$$\begin{aligned} \sup_{\|x\|_{L^2(S)} \leq 1} \|Hx\|_{L^2(T)}^2 &= \sup_{\|x\|_{L^2(S)} \leq 1} \int_{s \in S} \left[\int_{t \in T} k(t, s) x(s) ds \right]^2 ds \\ &\leq \sup_{\|x\|_{L^2(S)} \leq 1} \int_{t \in T} \left[\|k(t, \cdot)\|_{L^2(S)} \|x\|_{L^2(S)} dt \right]^2 ds = \|k\|_{HS}^2. \end{aligned}$$

Proposition 3.3 (Opérateur adjoint). Soit $H : \mathcal{X} \rightarrow \mathcal{Y}$ un opérateur linéaire tel que $\|H\|_{HS} \leq \infty$. Il existe une unique application linéaire $H^* : \mathcal{Y} \rightarrow \mathcal{X}$ telle que

$$\forall (x, y) \in \mathcal{X} \times \mathcal{Y}, \quad \langle Hx, y \rangle = \langle x, H^*y \rangle. \quad (3.14)$$

Elle est appelée *opérateur adjoint* à H et satisfait $\|H\|_{2 \rightarrow 2} = \|H^*\|_{2 \rightarrow 2}$.

Proposition 3.4 (Opérateur adjoint d'un opérateur intégral). Soit $H : \mathcal{X} \rightarrow \mathcal{Y}$ un opérateur linéaire intégral tel que $\|H\|_{HS} \leq \infty$. L'opérateur adjoint H^* est un opérateur intégral dans $\mathcal{L}(\mathcal{Y}, \mathcal{X})$ de noyaux k^* défini par $k^*(s, t) = k(t, s)$.

11. de façon plus générale un opérateur d'un Hilbert dans lui-même

3.4 Théorie spectrale

La diagonalisation des matrices semi-définies positives et la décomposition en valeurs singulières sont probablement parmi les outils les plus utilisés en ingénierie (pensez par exemple à l'analyse en composante principales). On peut en une certaine mesure, étendre ces concepts à des opérateurs définis sur des espaces fonctionnels plus riches. La théorie devient cependant rapidement difficile et requiert des mathématiques avancées. Je me contente donc de donner quelques résultats et définitions emblématiques ci-dessous.

Définition 3.7 (Spectre d'un opérateur borné). Soit $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ un espace de Banach et $H : \mathcal{X} \rightarrow \mathcal{X}$ un opérateur borné (i.e. $\|Bx\|_{\mathcal{X}} \leq C\|x\|_{\mathcal{X}}$ pour tout $x \in \mathcal{X}$ et une constante $C < +\infty$). Le spectre de H est l'ensemble des valeurs $\lambda \in \mathbb{C}$ telles que $H - \lambda \text{Id}$ n'est pas inversible. On le note $\sigma(H)$.

Malheureusement, la définition ci-dessus est souvent difficile techniquement à exploiter. Le spectre peut par exemple être continu (cas d'une convolution sur \mathbb{R}) ou discret (cas d'une convolution sur \mathbb{T}).

La situation s'améliore nettement lorsqu'on travaille sur des espaces de Hilbert avec des opérateurs de Hilbert-Schmidt. On rappelle en effet les deux résultats centraux suivants.

Théorème 3.2 (Théorème spectral sur les Hilbert). Soit \mathcal{X} un espace de Hilbert séparable et $H : \mathcal{X} \rightarrow \mathcal{X}$ un opérateur de Hilbert-Schmidt auto-adjoint : $\langle Hx, x' \rangle = \langle x, Hx' \rangle$ pour tout $x, x' \in \mathcal{X}$.

Alors il existe une base hilbertienne de vecteurs propres de H , i.e. une famille orthogonale (e_n) telle que $He_n = \lambda_n e_n$, avec des valeurs propres réelles $\lambda_n > 0$. De plus, on peut ordonner la série (λ_n) de façon décroissante.

Théorème 3.3 (Décomposition en valeurs singulières). Soit \mathcal{X} et \mathcal{Y} des espaces de Hilbert séparables et $H : \mathcal{X} \rightarrow \mathcal{Y}$ un opérateur de Hilbert-Schmidt. On note $R \in \mathbb{N} \cup \{+\infty\}$ son rang.

Alors il existe une base hilbertienne $(e_n)_{1 \leq n \leq R}$ de \mathcal{X} , une base hilbertienne $(f_m)_{1 \leq m \leq R}$ de \mathcal{Y} et une suite $(\sigma_n)_{1 \leq n \leq R}$ décroissante de \mathbb{R}_+^* , telle que

- $He_n = \sigma_n v_n$.
- $H^* v_m = \sigma_m e_m$.
- Pour tout $x \in \mathcal{X}$ et $y \in \mathcal{Y}$:

$$Hx = \sum_{n=0}^R \sigma_n \langle e_n, x \rangle f_n, \quad H^* y = \sum_{m=0}^R \sigma_m \langle f_m, y \rangle e_m \quad (3.15)$$

Exemple 3.1 (Convolution sur \mathbb{T}). Soit $h \in L^2(\mathbb{T})$ un noyau et $H : x \mapsto h \star x$. On peut montrer que H est un opérateur borné de $L^2(\mathbb{T})$ dans $L^2(\mathbb{T})$ puisque son noyau intégral k est de norme L^2 bornée.

Dans ce cas, l'opérateur est diagonalisé par la base de Fourier $e_n(s) = \exp(-isn)$ et les valeurs singulières sont données par les coefficients de la série de Fourier : $\sigma_n = (\mathcal{F}h)[n]$.

Les deux théorèmes précédents et l'exemple me semblent essentiels à connaître pour se forger une intuition sur les opérateurs. Il présentent tout de même de nombreuses faiblesses en pratique :

- Ces théorèmes assurent l'existence de deux bases (e_n) et (f_m) , mais en général, rien ne dit comment les construire. De même, rien ne dit comment calculer le spectre (σ_n) , alors qu'on a vu en introduction qu'il était essentiel pour mieux comprendre la difficulté de résolution d'un problème inverse.
- Beaucoup d'opérateurs ne sont pas de Hilbert-Schmidt (ou plus généralement compacts). Par exemple un opérateur de convolution sur $L^2(\mathbb{R})$ n'est pas un opérateur de Hilbert-Schmidt et on peut montrer qu'il n'est pas compact. Ainsi, on a vu en introduction que son spectre pouvait être *continu*.
- La convolution doit avoir des conditions de bord symétriques pour être diagonalisée par la série de Fourier (on a travaillé sur le tore), or ce modèle est en général irréaliste.

Une des théories les plus riches pour étendre les concepts semble être la notion de triplets de Gelfand. Elle est par exemple recommandée pour les ingénieurs par des mathématiciens tels que Hans Feichtinger, mais je n'ai pas réussi à la cerner suffisamment bien pour la restituer sous une forme compréhensible et semble poser des soucis à des amis que je juge pourtant forts compétents.

3.5 Discrétisation d'opérateurs

Dans ce paragraphe, nous illustrons plusieurs techniques distinctes de discrétisation d'un opérateur.

3.5.1 Calcul exact

En général il est impossible de calculer exactement la valeur $Hx(t)$ pour une fonction x arbitraire. C'est un problème d'intégration qui est souvent résolu par des formules de quadratures. Il existe cependant un cas important où c'est possible :

- on dispose d'une forme analytique des formes linéaires a_m (ou du noyau k) et on sait qu'elles sont continues et à support compact.
- le signal x est une mesure discrète de la forme $x = \sum_{p=1}^P \alpha_p \delta_{s_p}$.

Dans ce cas, on a simplement :

$$(Ax)[m] = \langle a_m, x \rangle = \sum_{p=1}^P \alpha_p a_m(s_p). \quad (3.16)$$

Ce cadre est utile lorsqu'on cherche des sources ponctuelles, comme en super-résolution par localisation de molécule unique.

3.5.2 Echantillonnage

Une technique fréquente de discrétisation lorsqu'on dispose du noyau intégral k consiste à l'échantillonner. Prenons par exemple le cas d'un noyau $k : [0, 1] \times [0, 1] \rightarrow \mathbb{R}$. On peut choisir un entier $N \in \mathbb{N}$ et construire une matrice $\mathbf{H} \in \mathbb{K}^{M \times N}$ définie pour tout coefficient (m, n) par

$$\mathbf{H}[m, n] = k(m/M, n/N).$$

Cette technique est très fréquente, et peut dans certains cas être précise (e.g. noyau à bande limitée et échantillonnage à la fréquence de Shannon). Cependant, elle mène

fréquemment à des erreurs de modélisation qui peuvent avoir des conséquences fâcheuses.

Elle me permet d'introduire une notion qui est centrale pour le calcul rapide. Supposons que k soit R fois différentiable. On peut alors choisir un point $z \in S$ de référence et effectuer un développement de Taylor :

$$k(t, s) \simeq \sum_{r=0}^{R-1} \frac{\partial_s^r k(t, z)}{r!} (s - z)^r \quad \text{avec } \partial_s^r k = \frac{\partial^r}{\partial s^r} k. \quad (3.17)$$

Ceci revient à décomposer k par un *tenseur de rang* R :

$$k(t, s) \simeq \sum_{r=0}^{R-1} \phi_r(t) \psi_r(s) \quad \text{avec } \phi_r(t) = \frac{\partial_s^r k(t, z)}{r!} \quad \text{et } \psi_r(s) = (s - z)^r. \quad (3.18)$$

Après discrétisation des fonctions ψ_r et ϕ_r , on peut construire une matrice \mathbf{H}_R de rang R définie par

$$\mathbf{H}_R = \sum_{r=1}^R \boldsymbol{\psi}_r \boldsymbol{\phi}_r^T \quad \text{avec } \boldsymbol{\psi}_r[m] = \psi_r(m/M), \boldsymbol{\phi}_r[n] = \phi_r(n/N). \quad (3.19)$$

Cette matrice est une bonne approximation de \mathbf{H} et on peut contrôler son erreur en utilisant des bornes sur le résidu intégral de Taylor. Un produit matrice-vecteur par \mathbf{H} coûte $O(R(M + N))$ opérations contre $O(MN)$ pour le calcul direct après discrétisation.

A retenir

Si le noyau intégral k est lisse, on peut effectuer obtenir des approximations de faibles rang de matrices en utilisant des séries de Taylor. Ceci permet parfois de réduire les coûts de calcul.

3.5.3 Le cadre hilbertien

Un intérêt significatif de travailler sur des espaces de Hilbert est de pouvoir décomposer le noyau de l'opérateur H dans une base orthogonale en produit tensoriel. En utilisant l'hypothèse (3.13), on peut en effet appliquer le théorème de Fubini et écrire :

$$\begin{aligned} Hx &= \sum_m \langle Hx, f_m \rangle f_m \\ &= \sum_m \left\langle \sum_n \alpha_n H e_n, f_m \right\rangle f_m \\ &= \sum_m \sum_n \alpha_n \langle H e_n, f_m \rangle f_m \\ &= \sum_m \sum_n \alpha_n \left(\int_t \left[\int_s k(t, s) e_n(s) dt \right] f_m(t) dt \right) f_m \\ &= \sum_m \sum_n \alpha_n \left[\int_t \int_s k(t, s) e_n(s) f_m(t) dt ds \right] f_m \\ &= \sum_m \sum_n \alpha_n \langle k, f_m \otimes e_n \rangle f_m, \end{aligned}$$

où le *produit tensoriel* $f_m \otimes e_n$ est défini pour tout (t, s) par

$$(f_m \otimes e_n)(t, s) \stackrel{\text{def}}{=} f_m(t)e_n(s),$$

et où le dernier produit scalaire est celui sur l'espace $L^2(T \times S)$. On voit donc que sous les hypothèses précédentes, l'action d'un opérateur H est entièrement décrite par la "matrice infinie" de coefficients

$$H_{m,n} = \langle He_n, f_m \rangle = \langle k, f_m \otimes e_n \rangle. \quad (3.20)$$

Notons en particulier que l'égalité de Parseval donne

$$\|H\|_{HS}^2 = \|k\|_{L^2(T \times S)}^2 = \sum_m \sum_n H_{m,n}^2. \quad (3.21)$$

Cette décomposition élémentaire a de nombreuses conséquences pour l'analyse numérique des problèmes inverses : i) on peut discrétiser de façon naturelle un opérateur en tronquant les doubles sommes ci-dessus, ii) l'étude des opérateurs sur des espaces de Hilbert se ramène à l'étude de matrices et iii) pour effectuer des calculs numériques rapides, on peut chercher des bases (e_n) et (f_m) telles que l'essentiel des coefficients $H_{m,n}$ sont négligeables.

Définissons par exemple un opérateur approché $\tilde{H}_{M,N} : \mathcal{X} \rightarrow \mathcal{Y}$ par

$$\tilde{H}_{M,N}x \stackrel{\text{def}}{=} \sum_{m=1}^M \sum_{n=1}^N H_{m,n} \langle x, e_n \rangle f_m, \quad (3.22)$$

on obtient immédiatement en utilisant l'identité de Parseval, le contrôle d'erreur suivant

$$\|\tilde{H}_{M,N} - H\|_{HS}^2 = \sum_{m>M} \sum_{n>N} H_{m,n}^2.$$

Une remarque importante est que la troncature de la série peut être vue comme une *approximation de rang MN* du noyau intégral k . Elle revient effectivement à approcher k par

$$\tilde{k}(x, y) = \sum_{m=1}^M \sum_{n=1}^N \lambda_{m,n} f_m(t) e_n(s) \quad \text{avec} \quad \lambda_{m,n} = \langle k, f_m \otimes e_n \rangle.$$

Notons que l'opérateur $H_{M,N}$ est codé avec uniquement $M \cdot N$ coefficients et peut donc être stocké sur un ordinateur. Dans beaucoup de travaux en imagerie que j'ai pu voir, la question de la discrétisation d'un opérateur est traitée rapidement : on se contente d'expériences dans un cadre discret-discret avec des matrices de taille finie. Je souhaite montrer ci-dessous que cette question est loin d'être anodine et peut avoir de grandes conséquences pour réduire les temps de calcul par exemple en utilisant des bases d'ondelettes orthogonales.

3.5.4 Méthodes de Galerkin

Une façon fréquente de discrétiser un opérateur, i.e. de se ramener du cadre continu à un cadre discret, consiste à approcher les éléments de \mathcal{X} par des éléments d'un sous-espace vectoriel $\mathcal{X}_N \subset \mathcal{X}$ défini par

$$\mathcal{X}_N = \text{span}(e_1, \dots, e_N).$$

Pour tout vecteur $x = \sum_{n=1}^N \alpha_n e_n \in \mathcal{X}_N$, l'action d'un opérateur intégral $A : \mathcal{X} \rightarrow \mathbb{R}^N$ revient à calculer

$$(Ax)[m] = \sum_{n=1}^M \alpha_n \langle a_m, e_n \rangle.$$

Ainsi, si on restreint x à \mathcal{X}_N , le produit Ax est entièrement déterminé par les coefficients matriciels

$$A_{m,n} = \langle a_m, e_n \rangle = \int_{s \in S} a_m(s) e_n(s) ds.$$

Ceux-ci peuvent être évalués par des formules de quadrature par exemple. On peut ainsi calculer l'action d'un opérateur intégral avec une excellente précision en se restreignant à des sous-espaces.

De plus, si l'espace \mathcal{X} contient des fonction suffisamment régulières, on peut contrôler l'erreur uniforme

$$\text{dist}(\mathcal{X}_N, \mathcal{X}) = \sup_{x \in \mathcal{X}, \|x\|_{\mathcal{X}} \leq 1} \inf_{x' \in \mathcal{X}_N} \|x' - x\|_{\mathcal{X}}, \quad (3.23)$$

avec des taux du type $\text{dist}(\mathcal{X}_N, \mathcal{X}) = O(N^{-\eta})$ pour un certain $\eta > 0$. Je renvoie le lecteur intéressé à des ouvrages comme [8] pour plus de détails.

3.6 Calcul matriciel rapide

Une fois qu'un opérateur H a été discrétisé en une matrice \mathbf{H} , par une technique ou une autre, se pose la question du calcul rapide. De nombreuses questions peuvent être posées :

- Comment calculer un produit matrice-vecteur $\mathbf{H}\mathbf{x}$?
- Comment calculer un produit matrice-vecteur $\mathbf{H}^*\mathbf{y}$?
- Comment calculer la solution d'un problème de type $\mathbf{H}^*\mathbf{H}\mathbf{x} = \mathbf{c}$?
- Comment calculer l'inverse ou la pseudo-inverse de \mathbf{H} ?
- Comment effectuer une décomposition LU de \mathbf{H} .

Dans ces notes, nous nous focaliserons sur les deux premiers problèmes. En effet dans beaucoup d'applications, il est seulement nécessaire de pouvoir calculer très efficacement les produits matrice-vecteur par \mathbf{H} et \mathbf{H}^* (e.g. méthodes d'optimisation). Le troisième point est souvent crucial aussi. On peut le résoudre avec des techniques de gradient conjugué linéaire si on sait évaluer les produits par \mathbf{H} et \mathbf{H}^* .

3.6.1 Le calcul naïf, ses limites et PyKeops.

Temps de calcul sur CPU Soit $\mathbf{H} \in \mathbb{K}^{M \times N}$ une matrice. Un produit matrice-vecteur avec \mathbf{H} et \mathbf{H}^* a une complexité en $O(MN)$. Prenons un exemple simple : supposons qu'on dispose d'une (petite) image de taille $N = 1000 \times 1000$ et que $M = N$ (e.g. une convolution). Un produit par \mathbf{H} coûte 10^{12} opérations (un téra). Avec un processeur à 1Ghz, un seul produit prend donc environ 1000 secondes¹², soit plus de 15'. En général, dans des méthodes itératives, il faut effectuer une centaine de produits, soit environ 1 journée pour effectuer une seule inversion! Pour peu qu'on se soit trompé de paramètre de régularisation, on imagine bien qu'une telle méthode est sinon inutilisable, au moins terriblement frustrante.

12. Un microprocesseur réalise typiquement 4FLOP par cycle d'horloge, mais si on travaille en double précision et qu'on réalise qu'il faut $4MN$ FLOP pour un produit.

Soucis de mémoire Un autre souci de taille apparaît : si la matrice est codée en double précision (soit 64 bits = 8 octets par coefficient), il faut stocker $8 \cdot 10^{12}$ octets, soit 8To. A moins d'avoir une très (très (très)) grosse mémoire vive, stocker la matrice est impossible.

Ainsi, jusqu'il y a 5 ans environ, j'aurais vivement sermonné un étudiant qui aurait eu le culot d'effectuer un calcul naïf. Les temps changent et aujourd'hui, je l'encouragerais à commencer par là, avec des outils adaptés.

PyKeops et les GPU Le calcul sur architecture GPU change en effet complètement la donne. Un calculateur GPU grand public (moins de 1kE), permet en effet de réaliser environ 10^{13} opérations par seconde (10 Tflops) en puissance de pointe. C'est monstrueux et ça résoud le problème des temps de calcul : dans l'exemple précédent, on passe de 15' par produit à 4 dixièmes de seconde en double précision !

Le souci de la mémoire est cependant critique puisque les meilleurs GPU atteignent péniblement 48Go aujourd'hui. Si on dispose d'une formule analytique des coefficients de \mathbf{H} par contre, 3 français (Benjamin Charlier, Jean Feydy et Joan Alexis Glaunès) ont récemment proposé un package Pytorch appelé PyKeops [3]. Celui-ci permet notamment d'effectuer des produits matrice-vecteur de façon très efficace. Il implémente ce qu'on appelle des réductions (sommes de termes et produits-scalaires notamment) de façon très efficace.

PyKeops permet notamment de calculer des opérations du type suivant :

$$K : \mathbb{R}^{M \times D} \times \mathbb{R}^{N \times D} \times \mathbb{R}^{N \times E} \rightarrow \mathbb{R}^{M \times E}$$

$$((t_m), (s_n), (x_n)) \mapsto (a_m) \quad \text{avec} \quad a_m = \sum_{n=1}^N k(t_m, s_n) x_n$$

où k est un noyau dont on connaît une expression analytique.

Un exemple concret Comme un exemple vaut mieux qu'un long discours, imaginons que l'on souhaite calculer un produit matrice-vecteur avec une matrice \mathbf{H} dont les termes généraux ont l'expression suivante :

$$\mathbf{H}[m, n] = \exp\left(-\frac{\|t_m - s_n\|^2}{2}\right) \quad (3.24)$$

où $(t_m) \in T$ et $(s_n) \in S$ sont un ensemble de points de \mathbb{R}^D . Si les points sont sur une grille cartésienne, \mathbf{H} devient une matrice de convolution, mais ici, on peut par exemple imaginer que les points sont hors grille. Dans ce cas, la matrice ci-dessus apparaît très fréquemment, par exemple dans le cadre des RKHS (Reproducible Kernel Hilbert Space), pour faire de l'interpolation de données dispersées. Voyons comment effectuer un produit matrice-vecteur en PyKeops.

```
import torch
from pykeops.torch import LazyTensor

M = 1000000
N = 2000000
D = 2 # Dimensions de s et t
E = 3 # Nombre d'images x a traiter

# On genere des positions et images aleatoires sur GPU
t = torch.randn(M, D).cuda()
```

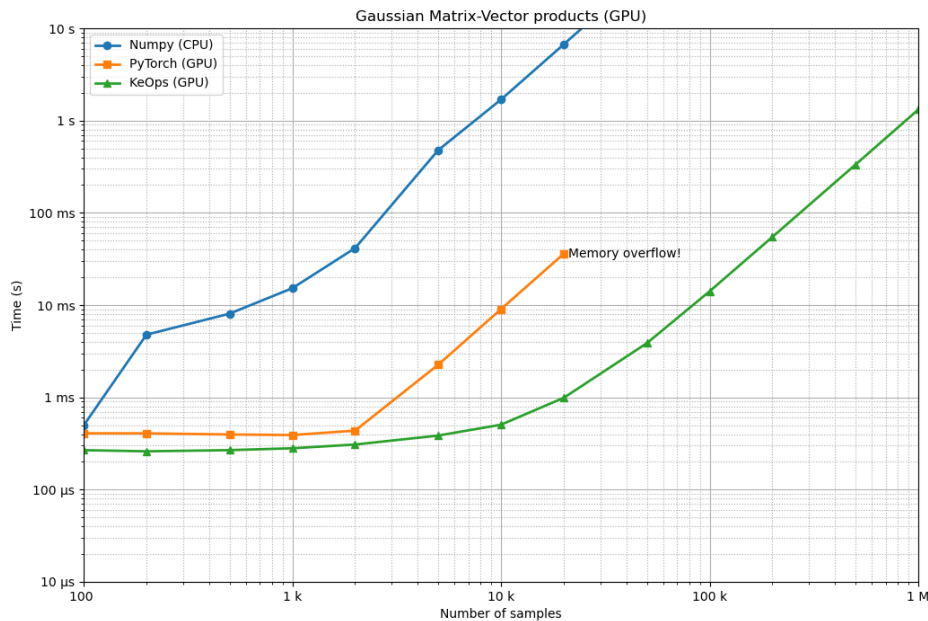


FIGURE 7 – Temps de calcul VS dimension. Ici on prend $M = N$. Voir le site [PyKeops](#).

```
s = torch.randn(N, D).cuda()
x = torch.randn(N, E).cuda()

# On transforme ca en LazyTensor (le coeur de PyKeops)
t_m = LazyTensor(t[:,None,:]) # (M,1,D)
s_n = LazyTensor(s[None,:,:]) # (1,N,D)

D_mn = ((t_m - s_n)**2).sum(dim=2) # Matrice symbolique (M,N,1) distance^2
K_mn = (- D_mn).exp() # Matrice symbolique (M,N,1) gaussienne
y = K_mn @ x # Produit matrice-vecteur
```

Pour cette matrice non structurée $2 \cdot 10^6 \times 10^6$, le produit matrice-vecteur prend moins de 2 secondes sur mon GPU!

Quand la légereté de l'être rappelle les scientifiques à la réflexion Heureusement, le minage des crypto-monnaies a causé une rupture de stocks des calculateurs GPU. Les temps nécessaires pour vous procurer une carte graphique raisonnable vous permettra donc de réfléchir un peu plus à des techniques qui font appel au raisonnement. Du moins, si vos directeurs de thèse n'ont pas eu la mauvaise idée de vous équiper de machines avec des calculateurs GPU récents!

3.6.2 Décomposition d'opérateurs dans des bases d'ondelettes orthogonales

Dans ce paragraphe, je vais présenter des idées proposées par Yves Meyer qui ont motivé la construction des bases d'ondelettes. La référence que j'ai utilisée avec P. Escande pour comprendre ces idées au départ est [18]. Dans [4], Albert Cohen

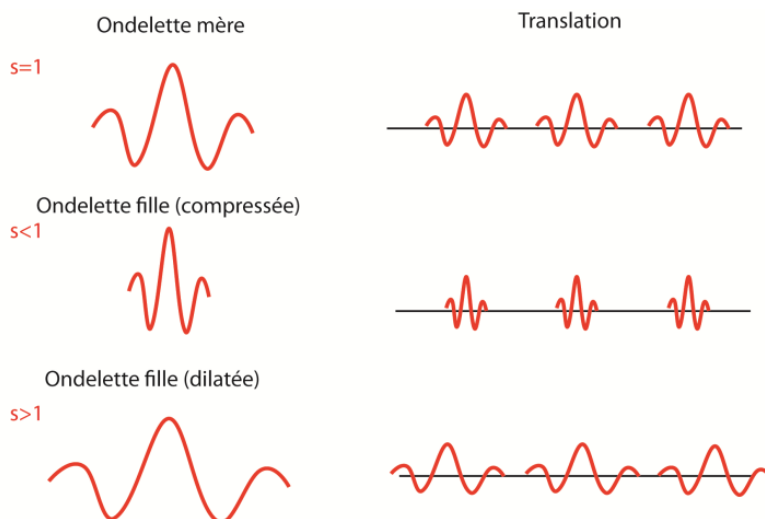


FIGURE 8 – Une ondelette mère et ses versions translatées et dilatées.

(étudiant de Yves Meyer) reprend ces idées sous une forme peut-être plus accessible, avec un intérêt particulier pour la résolution d'équations aux dérivées partielles.

Deux mots sur les bases d'ondelettes Je ne souhaite pas rentrer ici en détails dans la construction des bases d'ondelettes et renvoie le lecteur intéressé aux livres de références [17] pour une approche plus "ingénieur" et au livre [5] pour une introduction plus "mathématique".

Une ondelette (mère) sur $L^2([0, 1])$ est une fonction $\psi \in L^2([0, 1])$ de moyenne nulle. Une ondelette possède K moments nuls si

$$\int_{[0,1]} s^k \psi(s) ds = 0, \forall 0 \leq k < K. \quad (3.25)$$

Par exemple, l'ondelette de Haar

$$\psi(s) = \begin{cases} 1 & \text{si } 0 \leq s \leq 1/2 \\ -1 & \text{si } 1/2 \leq s \leq 1 \\ 0 & \text{sinon} \end{cases}$$

possède 1 moment nul. De façon plus générale, Ingrid Daubechies a construit une famille d'ondelettes à K moments nuls et de support compact de largeur minimale (l'ondelette de Haar coïncide avec celle de Daubechies pour $K = 1$).

Une ondelette peut être dilatée et translatée, ce qui génère une famille $(\psi_{j,n})_{j \in \mathbb{N}, n \in [0, 2^j - 1]}$ définie par

$$\psi_{j,n}(s) \stackrel{\text{def}}{=} 2^{j/2} \psi(2^j s - n).$$

Voir par exemple la figure 8.

Sous certaines conditions techniques, on peut montrer que la famille $(\psi_{j,n})_{m \leq 0, n \in [0, 2^j - 1]}$ complétée par une famille $(\phi_n)_{n \in [0, 2^j - 1]}$ (les ondelettes

père ou d'échelle) forme une base orthonormée de $L^2([0,1])$. Ainsi toute fonction $x \in L^2([0,1])$ s'écrit sous la forme

$$x = \sum_{j \geq 0, n \in \llbracket 0, 2^j - 1 \rrbracket} \langle x, \psi_{j,n} \rangle + \sum_{n \in \llbracket 0, 2^j - 1 \rrbracket} \langle x, \phi_n \rangle \phi_n. \quad (3.26)$$

Ces notations sont assez lourdes et on va préférer indiquer les ondelettes par un paramètre $\lambda = (j, n, t)$, où j est l'échelle, n le paramètre de translation et $t \in \{\text{père, mère}\}$, le type de l'ondelette (ψ ou ϕ). On note aussi Λ_J l'ensemble des indices λ jusqu'à l'échelle J . Ainsi une famille d'ondelettes orthogonales s'écrit sous la forme $(\psi_\lambda, \lambda \in \Lambda)$ avec $\Lambda = \Lambda_\infty$. Les ondelettes peuvent être étendue sur $[0,1]^D$ en utilisant des produits-tensoriels entre ondelettes. En 2D par exemple, on peut définir une base orthogonale en considérant tous les produits tensoriels de la forme $\psi_\lambda(s_1, s_2) = \psi_{\lambda'}(s_1)\psi_{\lambda''}(s_2)$ avec $\lambda = (\lambda', \lambda'')$. Cette définition correspond à ce qu'on appelle les ondelettes sous forme standard, mais il existe une autre forme plus populaire appelée non-standard que je ne souhaite pas décrire ici.

Etant donnée une échelle maximale $J \in \mathbb{N}$, on peut ainsi construire un opérateur d'analyse $\Psi_J^* : L^2([0,1]^D) \rightarrow \mathbb{R}^{N_J}$, où N_J est le nombre d'ondelettes jusqu'à l'échelle J . Cet opérateur associe à un signal x , ses coefficients en ondelettes dans \mathbb{R}^{N_J} :

$$\Psi_J^* x = (\langle x, \psi_\lambda \rangle, \lambda \in \Lambda_J).$$

L'opérateur de synthèse associé est $\Psi_J : \mathbb{R}^{N_J} \rightarrow L^2([0,1]^D)$ défini pour tout coefficient $\alpha \in \mathbb{R}^{N_J}$ par

$$\Psi_J \alpha = \sum_{\lambda \in \Lambda_J} \alpha_\lambda \psi_\lambda.$$

Pourquoi des moments nuls? Un oeil non exercé peut trouver la définition des moments nuls (3.25) étrange à première vue. La motivation provient de développements de Taylor locaux.

Définition 3.8 (Régularité Lipschitzienne). Une fonction $f \in L^2(S)$ est dite lipschitzienne d'ordre $\alpha > 0$ au point $s_0 \in S$ s'il existe une constante C_{s_0} et un polynôme de degré $\lfloor \alpha \rfloor$ tels que pour tout $s \in S$

$$|f(s) - P(s)| \leq C_{s_0} |s - s_0|^\alpha. \quad (3.27)$$

- La régularité Lipschitzienne de f en s_0 est le suprémum des α tel que (3.27) est vérifié.
- une fonction est dite uniformément lipschitzienne d'ordre α si (3.27) est vérifié pour tout $s_0 \in S$ où $C_{s_0} = C$ ne dépend pas de s_0 .

Ainsi une fonction discontinue en s_0 est d'exposant Lipschitz 0. Une fonction C^R sur un compact avec la dérivée R -ième bornée est d'exposant R . Si f est R fois continûment dérivable au voisinage de s_0 , alors on peut prendre pour polynôme $P(s) = \sum_{r=0}^R \frac{f^{(r)}(s_0)}{r!} (s - s_0)^r$. De plus, si

Si f est localement Lipschitzienne d'ordre k au voisinage de s_0 , on peut donc écrire

$$f(s) = P(s) + \epsilon(s) \quad \text{avec} \quad |\epsilon(s)| \leq |s - s_0|^\alpha \quad (3.28)$$

pour s proche de s_0 . Ainsi, si les ondelettes ont un petit support localisé en s_0 et possèdent α moments nuls, tous les produits scalaires $\langle f, \psi_\lambda \rangle$ vont être de faible amplitude et pourront être négligés. Ce phénomène est illustré sur la figure 9.

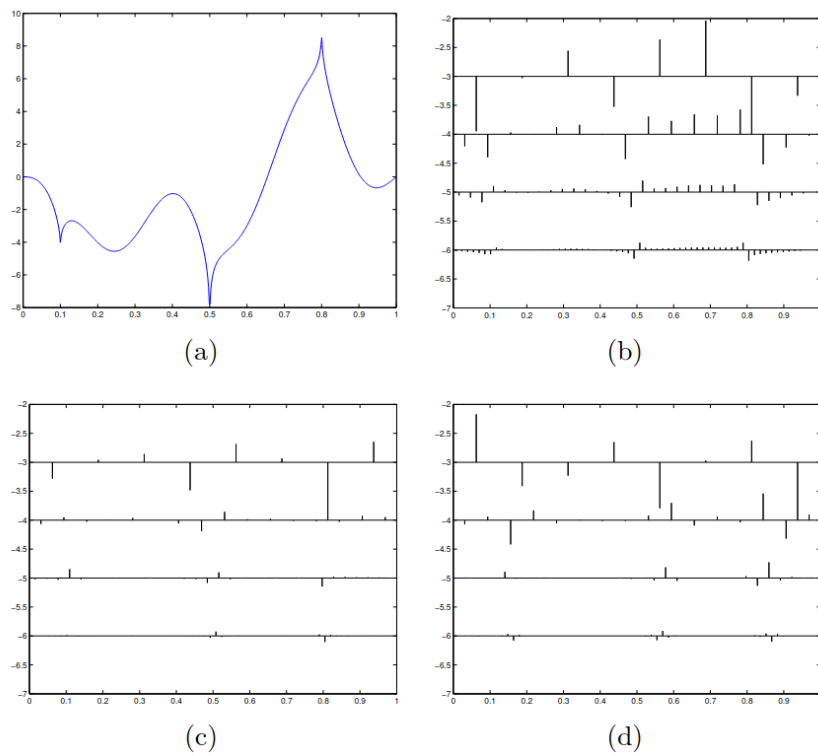


FIGURE 9 – Coefficients d’ondelettes d’une fonction polynomiale par morceaux. (b) Ondelette de Haar. (c) Ondelette de Daubechies, 2 moments nuls (d) Symmlets (4 moments nuls). Observer comme le nombre de coefficients nuls décroît avec le nombre de moments nuls du filtre.

Ondelettes et calcul numérique On peut alors construire la matrice de taille $N_J \times N_J$ définie par

$$\Theta = \Psi_J H \Psi_J^*. \quad (3.29)$$

La formule ci-dessus est identique à celle d'un changement de base pour des matrices, qui est un des standards de première année de licence/classe préparatoire. Notez que les coefficients $\Theta_{\lambda, \lambda'}$ de la matrice sont définis par

$$\Theta_{\lambda, \lambda'} = \langle H \psi_\lambda, \psi_{\lambda'} \rangle = \langle k, \psi_\lambda \otimes \psi_{\lambda'} \rangle. \quad (3.30)$$

La matrice Θ permet de coder l'opérateur H sur un ordinateur de façon efficace. Quitte à rester dans le domaine des coefficients d'ondelettes, effectuer un produit Hx avec un signal $x \in \text{span}(\psi_\lambda, \lambda \in \Lambda_J)$ est en effet équivalent à effectuer un produit matrice-vecteur par la matrice Θ . On peut alors s'intéresser aux propriétés de décroissance de la matrice Θ .

On peut alors énoncer le résultat suivant :

Théorème 3.4 (Complexité de calcul). Si le noyau intégral k appartient à $W^{\alpha, \infty}([0, 1]^D \times [0, 1]^D)$ avec $\alpha \in \mathbb{N}$ et si l'ondelette mère possède plus de $\lceil \alpha \rceil$ moments nuls, on peut construire une matrice $\tilde{\Theta}$ à partir de Θ telle que

$$\|\Theta - \tilde{\Theta}\|_{2 \rightarrow 2} \leq \eta$$

qui contient

$$O(N_J \log_2^2(N_J) \eta^{-D/\alpha})$$

coefficients non nuls.

Autrement dit, n'importe quel produit avec un opérateur intégral (stationnaire ou instationnaire) dont le noyau est suffisamment régulier peut être calculé avec une précision en $O(N \log_2^2(N))$! Pour des applications en imagerie (défloutage en particulier), je renvoie le lecteur intéressé aux articles [10, 12]. Nous avons notamment montré que des problèmes de défloutage pouvaient être accélérés de façon massive lorsqu'ils étaient appliqués dans le domaine des ondelettes.

Illustration Pour illustrer le résultat précédent, on peut considérer l'opérateur $H : L^2(\mathbb{T}) \rightarrow L^2(\mathbb{T})$ ¹³ dont le noyau k est défini par

$$k(t, s) \stackrel{\text{def}}{=} \frac{1}{\sigma(s)\sqrt{2\pi}} \exp\left(-\frac{(s-t)^2}{2\sigma(s)^2}\right) \quad \text{avec} \quad \sigma(s) \stackrel{\text{def}}{=} 4 + 10y. \quad (3.31)$$

Les réponses impulsionnelles de H dans ce cas sont simplement des gaussiennes dont l'écart-type varie linéairement sur l'intervalle $[0, 1]$. Après discrétisation, on peut afficher cet opérateur comme une matrice et effectuer la décomposition proposée précédemment. Les résultats sont illustrés sur la figure 10.

^{13.} un lecteur attentif remarquera une discontinuité du noyau k en 0 et 1, comme on travaille sur le tore

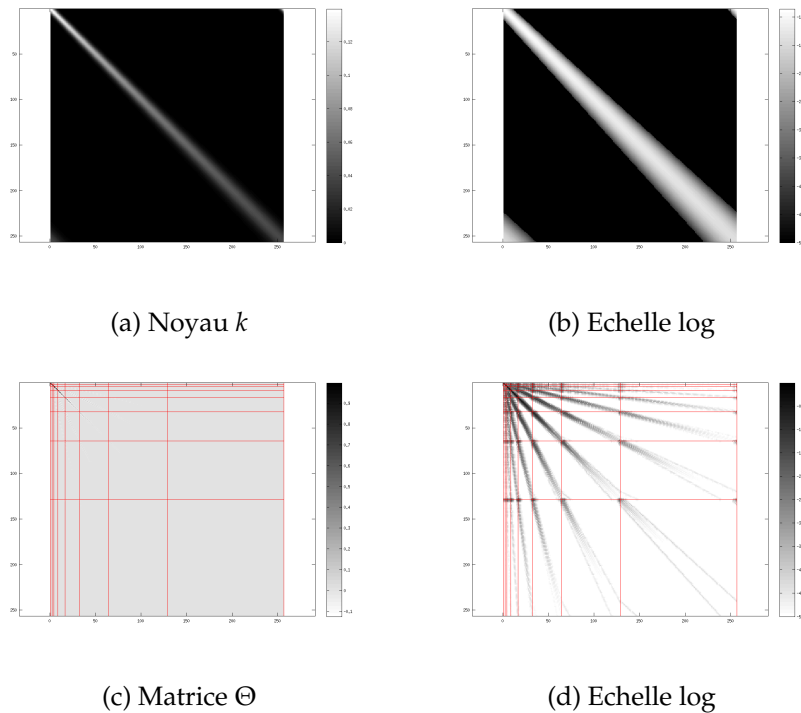


FIGURE 10 – Illustration d’une décomposition d’opérateur en ondelettes. Observer à quel point le nombre de grands coefficients de Θ est faible par rapport à celui de H .

3.6.3 Matrices \mathcal{H} (hiérarchiques) et méthodes multi-polaires

La décomposition précédente fait partie d’une famille plus générale de techniques multi-résolution pour coder efficacement des opérateurs intégraux. Une idée commune à toutes ces techniques est de décomposer une grande matrice en nombreuses sous-matrices et de remplacer l’action de chaque sous-matrice par celle d’une matrice de rang faible, obtenue par exemple avec un développement de Taylor du noyau. L’image 11 illustre une décomposition typique pour la discrétisation d’un opérateur intégral de $L^2([0, 1])$ dans $L^2([0, 1])$.

Je renvoie par exemple le lecteur intéressé au livre [13]. Bien qu’elles soient relativement peu connues ou utilisées dans le domaine signal/image, ces techniques sont loin d’être anecdotiques. Elles sont d’ailleurs considérées (de la même façon que la FFT) comme l’un des 10 algorithmes les plus importants du XXème siècle.

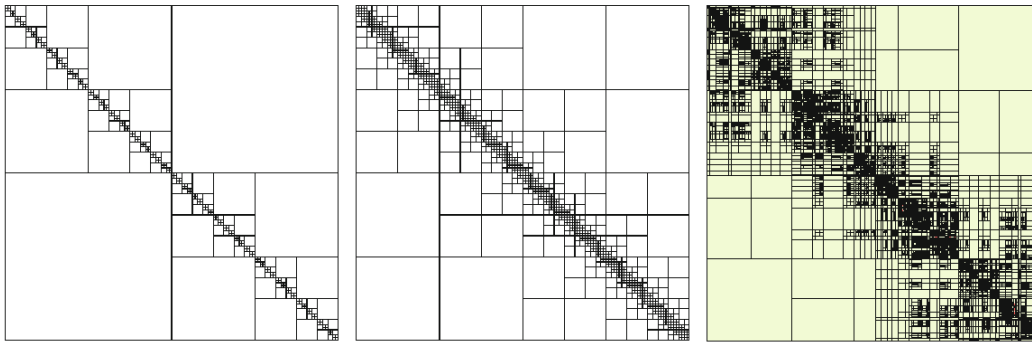


FIGURE 11 – Illustration d’une décomposition matricielle en sous-matrices. Tiré du livre [13].

A retenir

Aujourd’hui, on sait calculer des produits matrice-vecteur avec \mathbf{H} et \mathbf{H}^* , avec une bonne précision et une complexité numérique de l’ordre de $O((M + N)\text{polylog}(M + N))$, pour une très large collection de matrices structurées \mathbf{H} apparaissant dans les problèmes inverses.

Ces techniques me semblent encore trop peu populaires et je recommande vivement aux lecteurs ayant des problèmes inverses spécifiques de regarder cette littérature de près : il y a beaucoup à y gagner en termes de temps de calcul (et pour la recherche française, en termes de théorie de l’approximation).

3.6.4 Opérateurs de convolution et FFT.

Les matrices de convolution ont des propriétés structurelles qui permettent d’accélérer fortement la complexité numérique face à une implémentation naïve. Nous travaillons ici sur des signaux 1D, bien que les résultats s’étendent aisément en D dimension.

Soit $\mathbf{x} \in \mathbb{C}^N$ un vecteur et $\mathbf{h} \in \mathbb{C}^P$ un filtre discret.

Définition 3.9 (Convolution discrète). La convolution discrète entre \mathbf{x} et \mathbf{h} est un vecteur \mathbf{y} défini pour tout $m \in \mathbb{N}$ par

$$\mathbf{y}[m] = \sum_{n \in \mathbb{N}} \mathbf{x}[m - n] \mathbf{h}[n], \quad (3.32)$$

où la définition de \mathbf{h} et \mathbf{x} en dehors des intervalles $\llbracket 0, P - 1 \rrbracket$ et $\llbracket 0, N - 1 \rrbracket$ peut varier. Des choix populaires sont :

- Conditions de bord périodiques : $\mathbf{x}[n] = \mathbf{x}[n \bmod P]$.
- Conditions de bord nulles : $\mathbf{x}[n] = 0, n \notin \llbracket 0, N - 1 \rrbracket$.
- Conditions de bord libres : on étend le domaine de \mathbf{x} d’un facteur $\lceil P/2 \rceil$ à droite et à gauche et on laisse un algorithme (un régulariseur) combler l’information manquante. On ne garde que l’information pertinente sur $m \in \llbracket 0, N - 1 \rrbracket$.

Dans mon expérience, la dernière solution est très largement préférable pour des applications en défloutage par exemple. En effet, les autres choix sont simples, mais engendrent des erreurs de modélisation qui peuvent avoir des répercussions terribles sur la qualité de reconstruction.

Produit naïf La complexité numérique d'un produit de convolution discret sur $\llbracket 0, M-1 \rrbracket$ est $O(MP)$ (chaque coefficient $y[m]$ requiert P additions et multiplications). Dans les applications, M est typiquement égal ou proportionnel à N , et on obtient donc une complexité en $O(NP)$. Cette complexité est linéaire en N et tout à fait acceptable pour P petit. C'est d'ailleurs ce qui est fait par défaut dans des bibliothèques telles que TensorFlow, Keras ou PyTorch et donc dans les réseaux de neurones convolutionnels usuels.

Lorsque la taille du filtre \mathbf{h} est de l'ordre de N par contre (les filtres de l'optique ne sont pas à support borné), une complexité en $O(N^2)$ est inacceptable si les calculs se font sur CPU et peut être nettement réduite si les calculs se font sur GPU.

Convolutions et FFT Soit $\mathbf{F} \in \mathbb{C}^{N \times N}$ la matrice de la transformée de Fourier discrète définie pour tout $(m, n) \in \llbracket 0, N-1 \rrbracket^2$ par

$$\mathbf{F}[m, n] = \exp(-2i\pi mn/N). \quad (3.33)$$

La transformée inverse est donnée par

$$\mathbf{F}^{-1} = \frac{1}{N} \mathbf{F}^*. \quad (3.34)$$

On rappelle qu'un produit matrice-vecteur avec \mathbf{F} et \mathbf{F}^* peut être effectué $O(N \log(N))$ en utilisant la FFT (Fast Fourier Transform) qui dispose d'excellentes implémentations en CPU (e.g. FFTW) et en GPU sous le langage CUDA.

Un résultat central pour le calcul des convolutions est le suivant :

Théorème 3.5 (Diagonalisation des convolutions discrète). Soit $\mathbf{h} \in \mathbb{C}^N$ et $\mathbf{H} \in \mathbb{C}^{N \times N}$ une matrice de convolution discrète par \mathbf{h} avec des conditions de bord périodiques. Alors

$$\mathbf{H} = \mathbf{F}^{-1} \mathbf{\Sigma} \mathbf{F} \quad \text{avec} \quad \mathbf{\Sigma} = \text{diag}(\mathbf{F}\mathbf{h}). \quad (3.35)$$

Ce théorème est l'équivalent discret de la formule $\mathcal{F}(x \star y) = \mathcal{F}(x) \odot \mathcal{F}(y)$. Il indique que la transformée de Fourier diagonalise les opérateurs de convolution discrets. Sa conséquence en analyse numérique est la suivante : le coût d'une convolution par un filtre \mathbf{h} de taille inférieure à N peut être réalisé en $O(N \log(N))$ opérations en utilisant la FFT.

Un résultat moins connu permet d'effectuer des convolutions plus efficacement lorsque le filtre \mathbf{h} a une taille $P < N$. Dans ce cas, on peut appliquer des algorithmes appelés overlap-add ou overlap-save [19] qui permettent d'effectuer le produit de convolution en $O(N \log(P))$ opérations ! Ces algorithmes ont une complexité théorique toujours meilleure que les algorithmes naïfs $O(NP)$, tandis que la technique standard donnée par l'équation (3.35) n'est compétitive face aux algorithmes naïfs que si $\log(N) \lesssim P$. L'idée est de décomposer le domaine en petits intervalles disjoints et d'effectuer des convolutions circulaires pour chaque sous-domaine. Je

passer les détails sous silence ici. Ces mêmes techniques permettent aussi de gérer des conditions de bord arbitraires.

Toutes ces remarques peuvent être résumées comme suit.

Proposition 3.5 (Coût d'une convolution discrète). Une convolution discrète entre un filtre $\mathbf{u} \in \mathbb{C}^N$ supporté sur un intervalle de taille N^u et un signal $\mathbf{v} \in \mathbb{C}^N$ supporté sur un intervalle de taille N^v peut être effectuée en $O(N^v \log(N^u))$ opérations (en utilisant la transformée de Fourier rapide et une décomposition de domaine).

3.6.5 Convolution-produit (français) ou Product-convolution (anglais).

Commençons par une remarque sur les différences linguistiques entre le français et l'anglais. Si un opérateur H se décompose comme le produit $H = H_2 \circ H_1$, on va dire en français "l'opérateur $H_2 H_1$ ", tandis qu'en anglais, on dira "the operator $H_1 H_2$ ". La logique est la suivante : en anglais on décrit les opérateurs dans l'ordre où ils sont appliqués, tandis qu'en français, on les décrit dans l'ordre où ils sont écrits.

On commence par définir la structure de convolution-produit ci-dessous. Pour simplifier la discussion, on se place sur l'espace discret $\mathcal{X} = \mathbb{R}^N$ avec des conditions circulaires, mais on peut facilement étendre la discussion sur des espaces sur lesquels le produit de convolution et la multiplication entre fonctions est bien définie.

Définition 3.10 (Opérateur de convolution-produit). Soit $(\mathbf{u}_p)_{1 \leq p \leq P}$ et $(\mathbf{v}_p)_{1 \leq p \leq P}$ deux familles de vecteurs de \mathbb{R}^N . L'opérateur $\mathbf{H}_P : \mathbb{R}^N \rightarrow \mathbb{R}^N$ est un opérateur de convolution-produit d'ordre P s'il possède la structure :

$$\mathbf{H}_P \mathbf{x} = \sum_{p=1}^P \mathbf{u}_p \star \mathbf{v}_p \cdot \mathbf{x}, \quad (3.36)$$

où \odot est la multiplication terme à terme. Par la suite, les vecteurs \mathbf{v}_p seront nommés multiplicateurs et les vecteurs \mathbf{u}_p filtres.

Une propriété numérique intéressante des opérateurs de convolution-produit est la suivante :

Proposition 3.6 (Coût numérique d'un opérateur convolution-produit). Supposons que \mathbf{u}_p soit supporté sur un intervalle de taille N_p^u et que \mathbf{v}_p soit supporté sur un intervalle de taille N_p^v . Alors, une multiplication par \mathbf{H} peut être effectuée en $O\left(\sum_{p=1}^P N_p^v \log(N_p^u)\right)$ opérations et celle par \mathbf{H}^* en $O\left(\sum_{p=1}^P N_p^u \log(N_p^v)\right)$.

Démonstration. La multiplication par \mathbf{H}_P est une conséquence directe du paragraphe précédent. Pour \mathbf{H}_P^* , il suffit de se rendre compte que

$$\mathbf{H}_P \mathbf{x} = \sum_{p=1}^P \mathbf{v}_p \cdot \tilde{\mathbf{u}}_p \star \mathbf{x} \quad \text{où } \tilde{\mathbf{u}}_p = \mathbf{F}^2 \mathbf{u}_p \quad (3.37)$$

est le filtre \mathbf{u}_p symétrisé. □

Pour fixer un peu les idées, nous montrons un exemple pratique de décomposition en convolution-produit sur la figure 12.

Pour comprendre l'intérêt de la structure précédente, il me semble important de comprendre son lien avec la notion de réponse impulsionnelle variable.

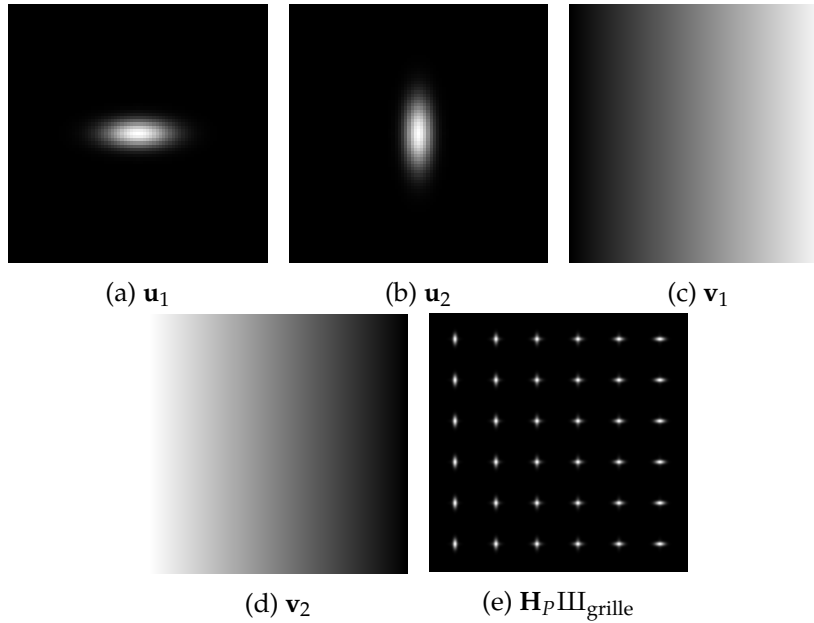


FIGURE 12 – Un exemple de décomposition en produit-convolution.

Définition 3.11 (RIV – réponse impulsionnelle variable). Soit $k \in L^2(\mathbb{R}^D \times \mathbb{R}^D)$ un noyau intégral¹⁴. On appelle *réponse impulsionnelle variable* qu'on abrègera en RIV la fonction définie pour tout (s, t) par

$$r(t, s) = k(s + t, s). \quad (3.38)$$

Pour illustrer la différence entre RIV et noyau intégral, on peut définir le noyau suivant sur $[0, 1] \times [0, 1]$:

$$k(t, s) \stackrel{\text{def}}{=} \exp\left(-\frac{(s-t)^2}{2\sigma(s)^2}\right) \quad \text{avec} \quad \sigma(s) \stackrel{\text{def}}{=} 0.01 + 0.02 \cos(\pi|t - 0.5|). \quad (3.39)$$

Le noyau (3.31) et la RIV correspondante sont affichées sur la figure 13. Une propriété essentielle qui est bien visible sur cet exemple est que la RIV semble de *rang numérique* plus faible que le noyau intégral. Cette propriété peut être vérifiée numériquement. Après discrétisation, on peut évaluer les valeurs singulières (le spectre) associé à k et r . Celui-ci est affiché en bas à droite sur la figure 13. On voit ici que le spectre de r décroît beaucoup plus rapidement que celui de k . En réalité, d'un point de vue numérique, le spectre de r s'annule au 8ème coefficient, bien que l'opérateur varie spatialement !

Que est le rapport entre les opérateurs de convolution-produit et la réponse impulsionnelle variable ? Il est capturé par le résultat suivant :

Proposition 3.7 (RIV, convolution-produit et rang faible). La RIV de l'opérateur $H : L^2(\mathbb{R}^D) \rightarrow L^2(\mathbb{R}^D)$ défini pour tout $x \in L^2(\mathbb{R}^D)$

$$Hx = \sum_{p=1}^P u_p \star v_p \odot x \quad (3.40)$$

14.

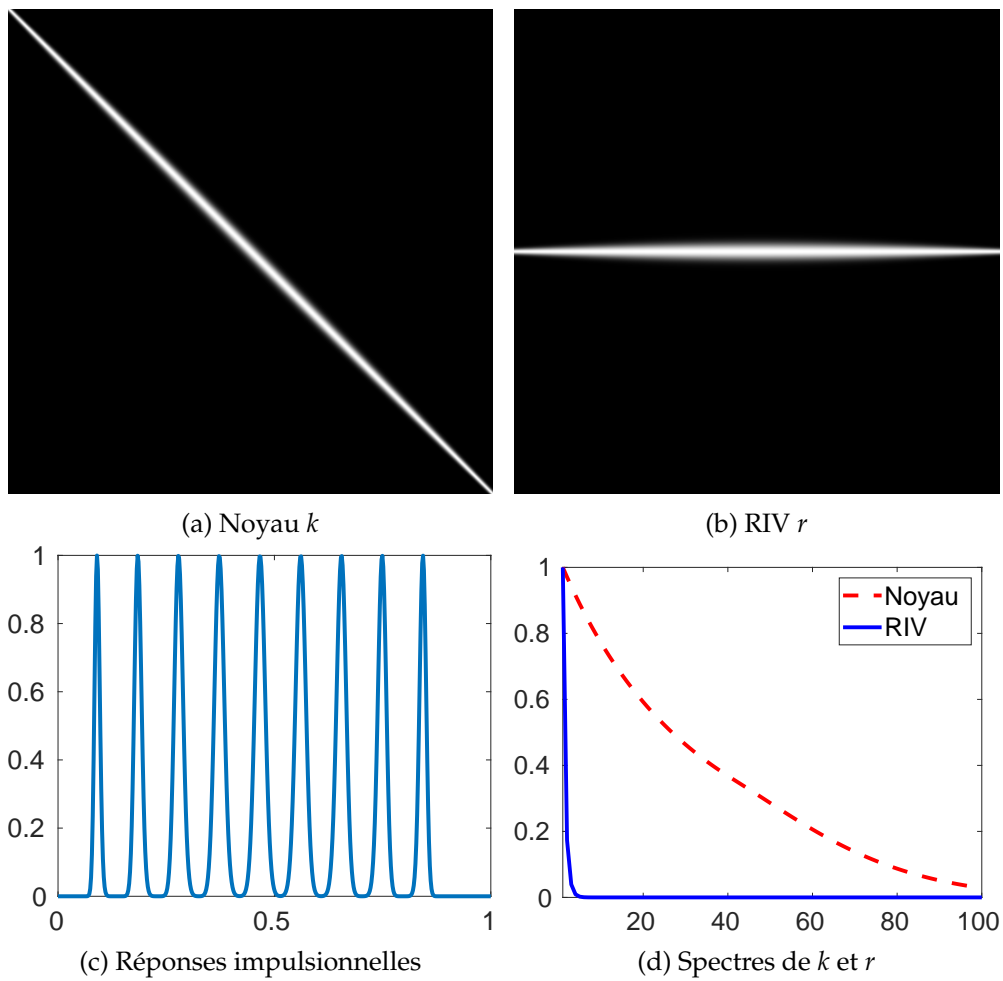


FIGURE 13 – En haut : Noyau intégral et sa RIV donnés par l'équation (3.39). En bas : quelques réponses impulsionnelles et spectres de k et r . Observer la vitesse de décroissance très rapide du spectre de r .

est donnée par le tenseur de rang P

$$r(t, s) = \sum_{p=1}^P u_p(t)v_p(s). \quad (3.41)$$

Ainsi, on voit qu'un opérateur de produit-convolution permet d'approcher efficacement les opérateurs dont la RIV est décomposable par un tenseur de rang faible, ou un tenseur de rang faible multi-échelle. On a vu dans les chapitres précédents le principe des matrices hiérarchiques et des décompositions en ondelettes. Toutes ces idées peuvent être reprises ici à la différence qu'on les applique à la RIV plutôt qu'au noyau intégral. La figure 14 illustre la différence entre les deux approches. On voit qu'une décomposition multi-échelle de la RIV semble plus adaptée à la structure "diagonale" des opérateurs qui sont essentiellement supportés sur une bande.

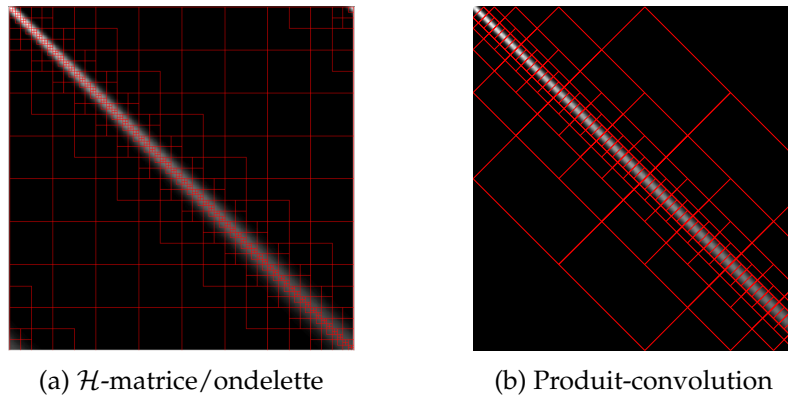


FIGURE 14 – Illustration d'une décomposition de domaine pour le noyau (à gauche) ou la SVIR (à droite) pour des méthodes multi-niveau. Observer que la décomposition de domaine appliquée à la SVIR donne ici des sous-matrices de rang plus faible et doit donc permettre de mieux compresser l'information, pour un coût similaire.

Les idées données ici sont appliquées de manière fréquente en imagerie optique sous des formes plus ou moins élémentaires.

Décompositions constantes par morceaux En optique, la façon la plus traditionnelle de décomposer un opérateur de flou variable spatialement, consiste à partitionner le domaine $S = [0, 1] \times [0, 1]$ en sous-domaines $S = \sqcup S_{i,j}$ avec $S_{i,j} = [(i-1)/I, i/I] \times [(j-1)/J, j/J]$. En choisissant $p = (i, j)$, on peut définir u_p et v_p par

$$u_p(s) = r\left(\left(\frac{i-0.5}{I}, \frac{j-0.5}{J}\right), s\right) \quad \text{et} \quad v_p(s) = \mathbb{1}_{S_{i,j}}. \quad (3.42)$$

Ceci revient à définir un opérateur dont la PSF est constante sur chaque domaine $S_{i,j}$ et égale à la PSF au milieu de $S_{i,j}$. J'ai vu cette technique utilisée de façon massive dans les articles de recherche, bien qu'elle ne me semble que peu recommandable ! En effet, des effets de bords inévitables apparaissent, ce qui change fortement le spectre de l'opérateur et peut résulter en des résultats de reconstruction bien mauvais.

Décomposition en valeur singulières Une autre possibilité qui pourrait sembler optimale consiste à construire la RIV discrète $\mathbf{R} \in \mathbb{R}^{N \times N}$ de $\mathbf{H} \in \mathbb{R}^{N \times N}$ puis à l'approcher par une matrice de rang P en calculant la SVD de \mathbf{R} :

$$\mathbf{R}_P \simeq \sum_{p=1}^P \mathbf{u}_p \otimes \mathbf{v}_p. \quad (3.43)$$

Cette technique est optimale au sens où pour un entier P donné, il n'existe pas de meilleure approximation de rang faible au sens de Hilbert-Schmidt. Elle possède deux défauts :

- Calculer les P premiers vecteurs singuliers de \mathbf{R} peut s'avérer très difficile lorsque N vaut quelques millions.
- La complexité numérique n'est en général pas optimale pour une erreur d'approximation donnée, car les filtres \mathbf{u} et \mathbf{v} n'ont pas en général un support de faible taille qui permettrait d'appliquer la proposition 3.5.

Décomposition en valeur singulière des PSF La matrice $\mathbf{R} = [\mathbf{r}_1, \dots, \mathbf{r}_N]$ contient les PSFs à toutes les positions dans l'espace.

Il est fréquent en imagerie que ces PSF soient exprimables dans une base commune de faible dimension (typiquement 3-10, penser par exemple à la décomposition en polynômes de Zernike dans le plan pupille). Ceci permet d'écrire que toute PSF \mathbf{r}_p se décompose sous la forme

$$\mathbf{r}_n \simeq \sum_{p=1}^P \langle \mathbf{r}_n, \mathbf{u}_p \rangle \mathbf{u}_p,$$

où la famille (\mathbf{u}_p) est orthogonale. En posant $\mathbf{U}_P = [\mathbf{u}_1, \dots, \mathbf{u}_P] \in \mathbb{R}^{N \times P}$ et $\mathbf{V}_P = \mathbf{R}^* \mathbf{U}_P = [\mathbf{v}_1, \dots, \mathbf{v}_P] \in \mathbb{R}^{N \times P}$, on obtient une décomposition naturelle de l'opérateur sous forme d'un tenseur de rang P : $\mathbf{R} \simeq \mathbf{U}_P \mathbf{V}_P^*$.

Cette idée présente de nombreux avantages : elle permet notamment d'estimer des opérateurs [2] ou des familles d'opérateurs [6] à partir de l'observation de réponses impulsionnelles dispersées. Elle correspond à la décomposition de la figure 12, à une orthogonalisation de Gram-Schmidt près des familles (\mathbf{u}) et (\mathbf{v}) .

Ondelettes et décompositions multi-échelles Finalement, les techniques les plus efficaces si le seul objectif est d'effectuer des produits rapides avec \mathbf{H} et \mathbf{H}^* sont celles qui reposent sur des décompositions multi-échelle de la RIV (\mathcal{H} -matrices ou ondelettes). Je ne détaille pas plus cette idée ici.

Pour plus de détails sur ces différentes approches, je recommande la lecture des deux articles qui - à ma connaissance - contiennent les idées les plus avancées sur ce domaine [7] (approche plus ingénieur) et [11] (approche plus mathématique).

4 Rappels d'algèbre linéaire

Nous précisons d'abord quelques notations.

- L'espace des matrices à m lignes et n colonnes est noté indifféremment $\mathbb{K}^{m \times n}$ ou $\mathcal{M}_{m,n}(\mathbb{K})$ avec $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} . Pour les matrices carrées de taille n , on notera indifféremment : $\mathbb{K}^{n \times n}$ ou $\mathcal{M}_{n,n}(\mathbb{K})$ ou $\mathcal{M}_n(\mathbb{K})$.
- Lorsqu'une matrice A possède n colonnes notées a_i , il est usuel d'écrire $A = (a_1, a_2, \dots, a_n)$ ou $A = [a_1, a_2, \dots, a_n]$. Suivant les disciplines, les crochets ou les parenthèses sont préférées. Dans les ouvrages de mathématiques, il semble que les parenthèses soient le symbole le plus souvent utilisé.
- Une matrice diagonale est une matrice *carrée* dont tous les éléments en dehors de la diagonale sont nuls. Lorsqu'on effectue une SVD d'une matrice rectangulaire, c'est donc un abus de langage que de dire " Σ est diagonale".
- On note la transposée d'une matrice sous la forme A^T .
- La matrice identité de dimension n sera notée I_n . Quelques-fois on se contentera de la notation I et on déduira sa dimension en fonction du contexte.

4.1 Image et noyau d'une matrice

Définition 4.1. L'image d'une matrice $A \in \mathbb{K}^{m \times n}$ est définie par :

$$\text{Im}(A) = \{Ax, x \in \mathbb{K}^n\} = \left\{ \sum_{i=1}^n x_i a_i, x \in \mathbb{K}^n \right\}.$$

Définition 4.2. Le rang d'une matrice est la dimension de son image. C'est donc le maximum du nombre de vecteurs colonnes indépendants.

Définition 4.3. Le noyau de A est le sous-espace vectoriel défini par :

$$\text{Ker}(A) = \{x \in \mathbb{K}^n, Ax = 0\}.$$

Théorème 4.1 (Théorème du rang). Si $A : E \rightarrow F$ est une application linéaire entre deux espaces vectoriels, le théorème du rang s'énonce ainsi :

$$\text{rang}(A) + \dim(\text{Ker}(A)) = \dim(E).$$

4.2 Déterminant

On renvoie aux cours des années 1 et 2 pour les définitions et la construction du déterminant comme une forme multi-linéaire alternée.

On rappelle simplement les propriétés suivantes :

- $\det : \mathbb{K}^{n \times n} \rightarrow \mathbb{K}$.
- $\det(AB) = \det(A)\det(B)$.
- $\det(A^T) = \det(A)$.
- A inversible $\Leftrightarrow \det(A) \neq 0$. Dans ce cas $\det(A^{-1}) = 1/\det(A)$

4.3 Valeurs propres et vecteurs propres

Définition 4.4. On appelle *valeur propre* d'une matrice $A \in \mathbb{K}^{n \times n}$ toute racine λ du polynôme caractéristique :

$$P_A(\lambda) = \det(A - \lambda I_n) = 0.$$

Définition 4.5. L'ensemble des valeurs propres de A est appelé *spectre* de A . Il est noté $\text{spec}(A)$.

Définition 4.6. On appelle *vecteur propre* associé à la valeur propre λ tout vecteur non nul qui satisfait :

$$Ax = \lambda x.$$

Définition 4.7. Une matrice $A \in \mathbb{K}^{n \times n}$ est dite *diagonalisable* si elle est semblable à une matrice diagonale. C'est-à-dire s'il existe une matrice $D \in \mathbb{K}^{n \times n}$ diagonale et une matrice $P \in \mathbb{K}^{n \times n}$ inversible telles que :

$$A = PDP^{-1}.$$

Dans ce cas, en écrivant $D = \text{diag}(\lambda_i)$, les valeurs λ_i représentent les valeurs propres de A . On peut choisir de prendre $P = (p_1, p_2, \dots, p_n)$ où les colonnes $p_i \in \mathbb{K}^n$ sont indépendantes et où p_i est un vecteur propre associé à λ_i .

4.4 Produit hermitien et produit scalaire

Soit E un espace vectoriel.

Définition 4.8. Une application $f : E \times E \rightarrow \mathbb{R}$ définit un *produit scalaire* si elle est :

— Semi-linéaire par rapport à la première variable. Ceci signifie que pour $a \in \mathbb{C}$, et $(x, y, z) \in E \times E \times E$,

$$f(ax + y, z) = af(x, z) + f(y, z).$$

— Linéaire par rapport à la seconde variable. Ceci signifie que pour $a \in \mathbb{C}$, et $(x, y, z) \in E \times E \times E$,

$$f(z, ax + y) = af(z, x) + f(z, y).$$

— Symétrique :

$$f(x, y) = f(y, x).$$

— Définie positive :

$$f(x, x) > 0, \forall x \neq 0.$$

Exemple 4.1. Le produit scalaire usuel (ou canonique) sur \mathbb{R}^n est défini par :

$$\langle x, y \rangle = \sum_{i=1}^n x_i y_i = x^T y = y^T x$$

Dans le cas où l'on travaille dans \mathbb{C} plutôt que \mathbb{R} , on utilise un produit scalaire hermitien.

Définition 4.9. Soit $f : E \times E \rightarrow \mathbb{C}$ une application. f est dite *hermitienne* si elle est :

- Semi-linéaire par rapport à la première variable. Ceci signifie que pour $a \in \mathbb{C}$, et $(x, y, z) \in E \times E \times E$,

$$f(ax + y, z) = \bar{a}f(x, z) + f(y, z).$$

- Linéaire par rapport à la seconde variable. Ceci signifie que pour $a \in \mathbb{C}$, et $(x, y, z) \in E \times E \times E$,

$$f(z, ax + y) = af(z, x) + f(z, y).$$

- A symétrie hermitienne :

$$f(x, y) = \bar{f}(y, x).$$

- Définie positive :

$$f(x, x) > 0, \forall x \neq 0.$$

Une telle fonction définit un produit scalaire hermitien (souvent appelé produit scalaire) et noté d'une des façons suivantes :

$$f(x, y) = \langle x, y \rangle = (x|y).$$

Exemple 4.2. Le produit scalaire hermitien usuel (ou canonique) sur \mathbb{C}^n est défini par :

$$\langle x, y \rangle = \sum_{i=1}^n \bar{x}_i y_i = \bar{x}^T y$$

ou

$$\langle x, y \rangle = \sum_{i=1}^n x_i \bar{y}_i = \bar{y}^T x$$

La notion de produit scalaire (réel dans \mathbb{R} et hermitien dans \mathbb{C}) permet de définir l'adjoint d'un opérateur linéaire.

Définition 4.10. Soit $A : E \rightarrow F$ un opérateur linéaire entre deux espaces de Hilbert munis de produits scalaires notés $\langle \cdot, \cdot \rangle_{E \times E}$ et $\langle \cdot, \cdot \rangle_{F \times F}$. L'adjoint de A noté A^* est défini comme l'unique opérateur linéaire (il existe toujours dans des Hilbert) tel que :

$$\langle Ax, y \rangle_{F \times F} = \langle x, A^* y \rangle_{E \times E}, \quad \forall (x, y) \in E \times F.$$

Exemple 4.3. Les exemples suivants sont fondamentaux :

- Si $A \in \mathcal{M}_{m,n}(\mathbb{R})$, son adjoint par rapport au produit scalaire usuel est $A^* = A^T$.
- Si $A \in \mathcal{M}_{m,n}(\mathbb{C})$, son adjoint par rapport au produit scalaire hermitien usuel est $A^* = \bar{A}^T$, la transposée de la conjuguée de A .

4.5 Isométries, matrices unitaires, matrices orthonormales et orthogonales

Les termes ci-dessus sont très semblables. Nous donnons leur définition précise.

Définition 4.11. Une *isométrie* $f : E \rightarrow F$ où E et F sont deux espaces métriques munis de normes $\| \cdot \|_E$ et $\| \cdot \|_F$, est une transformation qui conserve les distances : $\|f(x) - f(y)\|_F = \|x - y\|_E, \forall (x, y) \in E \times E$.

Définition 4.12. Une *matrice orthogonale* A est une matrice carrée réelle, dont les colonnes sont orthogonales deux à deux et de norme euclidienne égale à un. Ceci permet d'obtenir immédiatement les propriétés suivantes :

- $A^T A = A A^T = I$.
 - $A^{-1} = A^T$.
 - $\langle Ax, Ay \rangle = \langle x, y \rangle$, $\forall (x, y) \in \mathbb{R}^n \times \mathbb{R}^n$ et pour le produit scalaire usuel sur \mathbb{R}^n .
- Une telle matrice est aussi appelée souvent *matrice orthonormale*, car ses colonnes et lignes sont normées.

Définition 4.13. Une *matrice unitaire* A est une matrice carrée à coefficients complexes, dont les colonnes sont orthogonales deux à deux et de norme euclidienne égale à un. Ceci permet d'obtenir immédiatement les propriétés suivantes :

- $A^* A = A A^* = I$, où A^* est la transposée de la conjuguée complexe de A .
- $A^{-1} = A^*$.
- $\langle Ax, Ay \rangle = \langle x, y \rangle$, $\forall (x, y) \in \mathbb{C}^n \times \mathbb{C}^n$ et pour le produit scalaire hermitien usuel sur \mathbb{C}^n .

Remarque 4.1. Les matrices unitaires et orthogonales définissent toutes deux des applications linéaires qui conservent les distances. On les appelle donc quelques fois des isométries.

4.6 Normes matricielles

Il existe de très nombreuses normes matricielles. Bien que toutes soient équivalentes (car les matrices sont des objets de dimension finie), l'une ou l'autre définition doit être préférée suivant le problème à modéliser. Nous définissons les principales normes ci-dessous.

Définition 4.14. Soient E et F deux espaces vectoriels normés de normes $\|\cdot\|_E$ et $\|\cdot\|_F$ et $A : E \rightarrow F$ une application linéaire. On appelle norme induite de A et on note $\|A\|$ ou $\|A\|_{E,F}$ la quantité suivante :

$$\|A\| = \max_{\|x\|_E \leq 1} \|Ax\|_F.$$

Cette définition permet de créer des normes à l'envi sur les espaces de matrices. Nous rappelons uniquement la plus courante ci-dessous :

Définition 4.15. Soit $x \in \mathbb{K}^n$. Il est usuel de noter :

$$\|x\|_1 = \sum_{i=1}^n |x_i| \text{ Norme } l^1$$

$$\|x\|_2 = \sqrt{\sum_{i=1}^n |x_i|^2} \text{ Norme } l^2$$

$$\|x\|_\infty = \max_{1 \leq i \leq n} |x_i| \text{ Norme } l^\infty$$

Définition 4.16. La *norme spectrale* d'une matrice $A \in \mathcal{M}_{m,n}(\mathbb{K})$ est la norme induite par les normes canoniques sur \mathbb{K}^n et \mathbb{K}^m . Elle est définie par :

$$\|A\|_2 = \sup_{\|x\|_2 \leq 1} \|Ax\|_2.$$

Remarque 4.2. On a vu en cours que cette norme correspond à la plus grande valeur singulière de A . Il faut savoir de redémontrer.

Remarque 4.3. Cette norme est liée aux valeurs propres de A^*A . On note $\rho(A^*A)$ la plus grande valeur propre de A^*A . Alors :

$$\|A\|_2 = \sqrt{\rho(A^*A)}.$$

Définition 4.17. Soient A et B deux matrices de $\mathcal{M}_{m,n}(\mathbb{K})$. On peut définir :

$$\langle A, B \rangle_F = \text{trace}(B^*A) = \sum_{i,j} \bar{b}_{i,j} a_{i,j}.$$

Cette application définit un produit scalaire sur $\mathcal{M}_{m,n}(\mathbb{K})$. La norme associée à ce produit scalaire s'appelle *norme de Frobenius* :

$$\|A\|_F^2 = \text{trace}(A^*A) = \sum_{i,j} |a_{i,j}|^2.$$

Proposition 4.1. Soit $A \in \mathcal{M}_{m,n}$ et $B \in \mathcal{M}_{n,p}$. La norme de Frobenius possède les propriétés suivantes :

- Elle est multiplicative : $\|AB\|_F \leq \|A\|_F \|B\|_F$.
- Elle est équivalente à la norme spectrale :

$$\|A\|_2 \leq \|A\|_F \leq \sqrt{n} \|A\|_2$$

- $\|AB\|_F \leq \|A\|_2 \|B\|_F$.
- Elle est invariante par transformation unitaire. Si U et V sont deux matrices unitaires alors $\|UAV\|_F = \|A\|_F$.

4.7 Valeurs singulières et valeurs propres

Rappelons d'abord le résultat de décomposition en valeurs singulières.

Théorème 4.2. Pour toute matrice $A \in \mathcal{M}_{m,n}(\mathbb{K})$ de rang r , il existe :

- Une matrice unitaire $U \in \mathcal{M}_n(\mathbb{K})$.
- Une matrice unitaire $V \in \mathcal{M}_m(\mathbb{K})$.
- Une matrice diagonale $\Sigma \in \mathcal{M}_{m,n}(\mathbb{R})$ qui s'écrit $\begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}$ avec $D = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r)$.

telles que :

$$A = V\Sigma U^*.$$

Les valeurs $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$ sont appelées *valeurs singulières* de A .

Remarque 4.4. On peut écrire $A = V\Sigma U^*$ ou $A = U\Sigma V^*$ (le nom des matrices unitaires n'a pas d'importance). Par contre, il est préférable d'écrire la matrice unitaire à droite sous forme adjointe.

La décomposition en valeurs singulières est très générale, dans le sens où elle s'applique à toute matrice rectangulaire $m \times n$. La décomposition en valeurs propres, en revanche, ne fonctionne que pour certaines matrices carrées. Néanmoins, quand elles sont toutes les deux définies, elles sont liées.

Remarque 4.5. Soit M une matrice qui admet la décomposition $M = U\Sigma V^*$. Ainsi $M^*M = V|\Sigma|^2V^*$ et $MM^* = U|\Sigma|^2U^*$.

Les matrices M^*M et MM^* sont toutes deux semi-définies positives et admettent donc une décomposition en valeurs propres sous la forme $M^*M = FSF^*$ où F est unitaire et S diagonale, constituée des valeurs propres de M^*M . on voit ainsi que :

- Les valeurs singulières de M sont les racines des valeurs propres de M^*M ! Cette remarque permet aussi de calculer les valeurs propres.
- De même les valeurs propres de M^*M sont les carrés des valeurs singulières de M .
- En outre, les colonnes de U (vecteurs singuliers à gauche) sont vecteurs propres pour MM^* , et les colonnes de V (vecteurs singuliers à droite) sont vecteurs propres de M^*M .

Références

- [1] Gilles Aubert and Pierre Kornprobst. *Mathematical problems in image processing : partial differential equations and the calculus of variations*, volume 147. Springer Science & Business Media, 2006.
- [2] Jérémie Bigot, Paul Escande, and Pierre Weiss. Estimation of linear operators from scattered impulse responses. *Applied and Computational Harmonic Analysis*, 47(3) :730–758, 2019.
- [3] Benjamin Charlier, Jean Feydy, Joan Glaunès, François-David Collin, and Ghislain Durif. Kernel operations on the gpu, with autodiff, without memory overflows. *Journal of Machine Learning Research*, 22(74) :1–6, 2021.
- [4] Albert Cohen. *Numerical analysis of wavelet methods*. Elsevier, 2003.
- [5] Ingrid Daubechies. *Ten lectures on wavelets*. SIAM, 1992.
- [6] Valentin Debarnot, Paul Escande, Thomas Mangeat, and Pierre Weiss. Learning low-dimensional models of microscopes. *IEEE Transactions on Computational Imaging*, 2020.
- [7] Loïc Denis, Eric Thiébaud, Ferréol Soulez, Jean-Marie Becker, and Rahul Mourya. Fast approximations of shift-variant blur. *International Journal of Computer Vision*, 115(3) :253–278, 2015.
- [8] Ronald A DeVore and George G Lorentz. *Constructive approximation*, volume 303. Springer Science & Business Media, 1993.
- [9] Ivar Ekeland and Roger Temam. *Convex analysis and variational problems*. SIAM, 1999.
- [10] Paul Escande and Pierre Weiss. Sparse wavelet representations of spatially varying blurring operators. *SIAM Journal on Imaging Sciences*, 8(4) :2976–3014, 2015.
- [11] Paul Escande and Pierre Weiss. Approximation of integral operators using product-convolution expansions. *Journal of Mathematical Imaging and Vision*, 58(3) :333–348, 2017.
- [12] Paul Escande and Pierre Weiss. Fast wavelet decomposition of linear operators through product-convolution expansions. *IMA journal of numerical analysis*, 2020.

- [13] Wolfgang Hackbusch. *Hierarchical matrices : algorithms and analysis*, volume 49. Springer, 2015.
- [14] Lars Hörmander. *The analysis of linear partial differential operators III : Pseudo-differential operators*. Springer Science & Business Media, 2007.
- [15] Alex Iosevich and Elijah Liflyand. *Decay of the Fourier Transform : Analytic and Geometric Aspects*. Springer, 2014.
- [16] Yitzhak Katznelson. *An introduction to harmonic analysis*. Cambridge University Press, 2004.
- [17] Stéphane Mallat. *A wavelet tour of signal processing*. Elsevier, 1999.
- [18] Yves Meyer. *Wavelets and Operators : Volume 1*. Number 37. Cambridge university press, 1992.
- [19] William D Stanley, Gary R Dougherty, Ray Dougherty, and H Saunders. *Digital signal processing*. 1988.