

Olga Holtz

Olga Holtz est une mathématicienne appliquée travaillant dans le domaine de l'analyse numérique, de l'algèbre linéaire numérique et de la théorie de l'approximation. Comme son nom l'indique, l'analyse numérique est la discipline qui étudie les solutions numériques aux problèmes d'analyse. Le but est de trouver des solutions approchées précises à des problèmes qui ne peuvent pas être résolus autrement. Nous avons déjà mentionné la méthode de Newton-Raphson comme exemple. Une autre branche importante de l'analyse numérique est l'intégration numérique, où l'on cherche à utiliser une méthode numérique pour évaluer une intégrale définie qui ne peut pas être calculée en termes de fonctions élémentaires (la plupart des intégrales appartiennent à cette catégorie). Commençons par une intégrale

$$\int_a^b f(x)dx.$$

Supposons que l'intégrande puisse être évaluée en des points équidistants et que l'intégrale puisse être approchée par une formule de la forme

$$\int_a^b f(x)dx \approx \frac{b-a}{N} \sum_{k=1}^N c_k f(a+kh).$$

Une formule de ce type est connue sous le nom de formule de Newton-Cotes. Par exemple, posons $N = 2$.

$$\int_a^b f(a+y)dx \approx \frac{b-a}{N} \sum_{k=1}^N c_k (f(a+kh)).$$

Utiliser les développements de Taylor.

$$f(a+kh) = f(a) + khf'(a) + \frac{1}{2}k^2h^2f''(a) + O(h^3),$$

$$f(a+y) = f(a) + \frac{1}{2}hf'(a) - \frac{1}{2}h^2f''(a) + O(h^3),$$

où h est le pas $(b-a)/N$. Ceci nous donne une représentation de Taylor qui doit être équivalente à la fonction originale sommée :

$$f(a) + h \left(\frac{3}{2}(f'(a) + hf''(a)) \right) + O(h^3).$$

Par comparaison des coefficients, nous avons $c_1 = c_2 = 3/2$. Ceci nous donne une formule de Newton-Cotes connue sous le nom de méthode des trapèzes :

$$\int_a^b f(x)dx \approx \frac{b-a}{3} \left(\frac{3}{2}(f_1 + f_2) \right)$$

Traduction de la biographie extraite de l'article arXiv <https://arxiv.org/pdf/1910.01730> *Twenty female mathematicians* de Hollis Williams.

Traduction et transcription en L^AT_EX : Denise Vella-Chemla, assistée d'outils de lecture d'image Google Lens et de traduction Google translate.

On peut utiliser un ordinateur pour trouver le troisième coefficient. Ceci nous donne une formule de Newton-Cotes connue sous le nom de méthode de Milne :

$$\int_a^b f(x)dx \approx \frac{b-a}{4} \left(\frac{4}{3}(2f_1 - f_2 + f_3) \right).$$

On peut continuer ainsi pour obtenir d'autres formules de Newton-Cotes.

L'idée d'approximer une intégrale par une somme qui lui est étroitement liée est courante en analyse. La formule d'Euler-Maclaurin, par exemple, donne la différence entre l'intégrale et la somme. Dans certains cas, on peut prendre une fonction définie par une intégrale et l'écrire sous forme de développement en série. Par exemple, commençons par la fonction suivante :

$$f(x) = \int_0^1 \frac{\sqrt{t}}{(1-xt^2)^\beta} dt,$$

où $\beta > 1$. On sait que

$$(1-z)^{-\beta} = \frac{1}{\Gamma(\beta)} \sum_{k=0}^{\infty} \frac{\Gamma(k+\beta)z^k}{k!},$$

où Γ est la fonction gamma. Cette fonction est essentiellement une généralisation de la fonction factorielle à l'analyse complexe. Pendant un certain temps, il a été suggéré que la caractéristique d'Euler $\chi(\text{Out}(F))$ pourrait être liée à la théorie des nombres. Borinsky a prouvé que c'est le cas en utilisant la fonction gamma. En particulier :

$$\bar{\chi}(\text{Out}(F_{n+1})) = -\frac{\Gamma(n+\frac{1}{2})}{n \log^2 n} + O\left(\frac{\Gamma(n+\frac{1}{2})}{n \log^4 n}\right).$$

$\chi(\text{Out}(F_{n+1}))$ croît à un rythme plus rapide qu'exponentiel, car le terme d'erreur n'est pas suffisant pour annuler la croissance. Ceci résout une conjecture de Vogtmann selon laquelle on peut prendre la caractéristique d'Euler des groupes de classes d'applications pour obtenir une relation avec les nombres de Bernoulli. Rappelons à partir de la définition de la fonction gamma que

$$\begin{aligned} \Gamma\left(n+\frac{1}{2}\right) &= \frac{2n!}{4n!} \sqrt{\pi}, \\ &= \frac{(2n-1)!!}{2^n} \sqrt{\pi}. \end{aligned}$$

Ceci implique l'identité intéressante :

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}.$$

Par comparaison avec la série ci-dessus, nous avons

$$\begin{aligned}
f(x) &= \int_0^1 dt \frac{\sqrt{t}}{\Gamma(\beta)} \sum_{k=0}^{\infty} \frac{\Gamma(k+\beta)}{k!} x^k t^{2k}, \\
&= \frac{1}{\Gamma(\beta)} \sum_{k=0}^{\infty} \frac{\Gamma(k+\beta)}{k!} x^k \int_0^1 dt t^{2k+\frac{1}{2}}, \\
&= \frac{1}{\Gamma(\beta)} \sum_{k=0}^{\infty} \frac{\Gamma(k+\beta) x^k}{k! (2k+\frac{3}{2})} \\
&= \frac{2}{\Gamma(\beta)} \sum_{k=0}^{\infty} \frac{\Gamma(k+\beta)}{k! (4k+3)} x^k.
\end{aligned}$$

Nous aimerions maintenant trouver le rayon de convergence de la série. Cela peut être fait très facilement en trouvant la limite du rapport des coefficients :

$$\begin{aligned}
u_k &= \frac{\Gamma(k+\beta)}{k! (4k+3)}, \\
\frac{u_k}{u_{k+1}} &= \frac{\Gamma(k+\beta)}{k! (4k+3)} \frac{(k+1)! (4k+4)}{\Gamma(k+\beta+1)}, \\
&= \frac{(k+1)(4k+4)}{4k+3} \frac{\Gamma(k+\beta)}{(k+\beta)\Gamma(k+\beta)}, \\
&= \frac{(k+1)(4k+4)}{(k+\beta)(4k+3)}.
\end{aligned}$$

Ce rapport tend vers 1 lorsque k tend vers l'infini, donc le rayon de convergence est 1. Nous prenons maintenant le rapport du coefficient avec le coefficient précédent et effectuons les calculs algébriques ci-dessous :

$$\begin{aligned}
\frac{u_k}{u_{k-1}} &= \frac{\Gamma(k-\beta)}{k! (2k+\frac{1}{2}+1)} \frac{(k-1)! (2k+\frac{1}{2}-2)}{\Gamma(k+\beta-1)}, \\
&= \frac{(k+\beta-1)(2k+\frac{1}{2}-2)}{k(2k+\frac{1}{2}+1)}, \\
&= \left(1 + \frac{\beta-1}{k}\right) \left(1 - \frac{3}{4k}\right) \left(1 + \frac{3}{4k}\right)^{-1} \\
&= \left(1 - \frac{3}{4k} + \frac{\beta-1}{k} + O(k^{-2})\right) \left(1 - \frac{3}{4k} + O(k^{-2})\right)
\end{aligned}$$

C'est-à-dire que

$$\begin{aligned}\frac{u_k}{u_{k-1}} &= 1 - \frac{3}{4k} + \frac{\beta^{-1}}{k} - \frac{3}{4k} + O(k^{-2}), \\ &= 1 + \frac{4(\beta - 1) - 6}{4k} + O(k^{-2}), \\ &= 1 + \frac{4\beta - 10}{4k} + O(k^{-2}), \\ &= 1 - \frac{\frac{10}{4} - \beta}{k} + O(k^{-2}).\end{aligned}$$

En comparant avec un résultat de Richards, on constate qu'il existe une singularité d'ordre $\alpha = 1 - \beta$, puisque¹

$$f(x) \sim A(1 - x)^{1-\beta},$$

pour une certaine constante A et pour x proche du rayon de convergence.

Une autre méthode courante d'intégration numérique est la quadrature de Gauss. On approxime une intégrale définie représentant une fonction comme une somme pondérée sur N points :

$$\int_{-1}^1 w(x)f(x)dx \approx \sum_{k=0}^N c_k f(x_k),$$

où $w(x)$ est une fonction de pondération. La fonction de pondération est souvent la fonction unité, mais pas toujours. *Attention* : les commandes de quadrature de base des programmes de calcul formel ne recherchent généralement pas les singularités, car elles supposent probablement une fonction de pondération unitaire. Comme nous le savons, la notation du produit scalaire désigne

$$(f, g) = \int_a^b w(x)f(x)g(x)dx.$$

Si nous considérons les fonctions de cette intégrale comme l'ensemble des polynômes orthogonaux $\phi_k(x)$ tels que le terme dominant de ϕ_n est x^n , nous pouvons supposer qu'ils satisfont une condition d'orthogonalité.

$$(\phi_n, \phi_m) = h_n \delta_{nm}.$$

En écrivant cela explicitement, on pourrait choisir la fonction de pondération de telle sorte que l'on ait

$$\int_0^1 (1 - x)^{3/2} \phi_n(x) \phi_m(x) dx = h_n \delta_{nm}.$$

Ces polynômes peuvent être trouvés en utilisant l'orthogonalisation de Gram-Schmidt. Par exemple,

1. Derek Richards, *Advanced Mathematical Methods with Maple* (Cambridge : Cambridge University Press, 2001).

les premiers polynômes seraient

$$\phi_0 = 1,$$

$$\phi_1 = x - \frac{2}{7},$$

$$\phi_2 = x^2 - \frac{8}{11}x + \frac{8}{99},$$

et ainsi de suite. Les coefficients correspondants du delta de Kronecker sont :

$$h_0 = \frac{2}{5}, \quad h_1 = \frac{8}{441}, \quad h_2 = \frac{128}{127.413}.$$

Supposons maintenant que nous souhaitions approximer l'intégrale définie suivante :

$$\int_0^1 (1-x)^{3/2} f(x) dx.$$

Écrire

$$f(x) = \sum_{j=0}^{\infty} a_j \phi_j(x),$$

où

$$a_j = \frac{1}{h_j} \int_0^1 w(x) \phi_j(x) f(x) dx.$$

On voit alors la similitude avec les coefficients de Fourier. Choisissons les x_k comme étant les N zéros du polynôme $\phi_N(x)$ et les c_k pour résoudre les équations linéaires

$$\sum_{j=1}^N c_j = h_0, \quad \sum_{j=1}^N c_j p_k(x_j),$$

pour k jusqu'à $N - 1$.

Choisissons $N = 4$, alors nous avons une approximation

$$\int_0^1 (1-x)^{3/2} f(x) dx \approx \sum_{k=1}^4 c_k f(x_k),$$

où les nombres x_k sont les zéros du polynôme orthogonal ϕ_4 du quatrième ordre, et les coefficients sont donnés par un tuple comme suit

$$c = [0.122, 0.169, 0.092, 0.017].$$

Cette quadrature peut être implémentée explicitement avec un ordinateur pour obtenir des valeurs numériques pour certaines intégrales définies. Par exemple,

$$\int_0^1 (1-x)^{3/2} \cos(\pi \sin x) dx \approx 0.221.$$

La méthode de quadrature de Gauss peut être étendue pour permettre des approximations d'intégrales de la forme

$$\int_{-1}^1 \frac{f(x)}{\sqrt{1-x^2}} dx.$$

Dans ce cas, la quadrature est la quadrature de Tchebychev-Gauss. Si nous prenons la fonction de pondération e^{-x^2} , nous avons la quadrature de Gauss-Hermite pour évaluer les intégrales de la forme

$$\int_{-\infty}^{\infty} e^{-x^2} f(x) dx.$$

Notons que cela permet une intégration sur un intervalle infini.

Olga Holtz a travaillé en particulier sur les polynômes de matrices orthogonales, étendant certains résultats classiques de la théorie des polynômes orthogonaux sur le cercle unité au cas matriciel.

Olga Holtz a prouvé ce que l'on peut appeler un théorème de Szegő matriciel : pour toute mesure de probabilité matricielle $\sigma \in P_\ell(\mathbb{T})$ et tout entier naturel n , on a

$$\int_{\mathbb{T}} \text{Tr} \log \sigma' \frac{d\theta}{2\pi} \leq \text{Tr} \log \beta_n = \log \prod_{k=0}^{n-1} \det(1 - \alpha_k^* \alpha_k).$$

En corollaire, on a que si σ est une mesure de Szegő, alors

$$\int_{\mathbb{T}} \text{Tr} \log \sigma' \frac{d\theta}{2\pi} \leq \inf_n \text{Tr} \log \beta_n \leq -\sup_n \log \|\beta_n^{-1}\| \leq 0.$$

En utilisant ces techniques, Olga Holtz a obtenu une preuve élémentaire de la formule de distance de Helson-Lowdenslager, qui stipule que pour tout $\sigma \in P_\ell(T)$,

$$\exp \int_{\mathbb{T}} \frac{1}{\ell} \text{Tr} \log \sigma' \frac{d\theta}{2\pi} = \inf_{A,P} \int_{\mathbb{T}} \frac{1}{\ell} \text{Tr}((A+P)^* d\sigma(A+P)),$$

où A parcourt toutes les matrices de déterminant unitaire et P parcourt tous les polynômes trigonométriques de la forme ²

$$P(e^{i\theta}) = \sum_{k>0} A_k e^{ik\theta}.$$

La plupart d'entre nous connaissent l'orthogonalisation de Gram-Schmidt. L'algèbre linéaire numérique plus avancée considère des méthodes itératives adaptées aux grands systèmes creux. Nous avons mentionné précédemment les méthodes de Gauss-Seidel et du gradient conjugué, et indiqué que la méthode de Gauss-Seidel n'est pas aussi efficace que la méthode du gradient conjugué. En fait, cela peut être quantifié en incluant une comparaison entre les deux méthodes en termes de temps de calcul nécessaire pour résoudre une matrice définie positive $n \times n$ générée aléatoirement

2. Maxim Derevyagin, Olga Holtz, Segey Krushchev et Mikhail Tyaglov, "Théorème de Szegő pour les polynômes orthogonaux matriciels", Journal of Approximation Theory, 164(9), 1238-1261 (2011).

avec une densité de 0,1. La comparaison porte sur le gradient conjugué contre Gauss-Seidel avec deux préconditionneurs différents³.

n	Gauss-Seidel	Gauss-Seidel (RCM)	Gauss-Seidel (MD)	CG
100	0.007928804	0.007056294	0.006839876	0.002512009
500	0.502849338	0.405763259	0.303914523	0.002419189
1000	3.219242909	3.281714839	2.618037809	0.004361086
3000	116.8587276	105.7410554	99.96394861	0.044996817

TABLEAU 2 : Efficacité de la méthode du gradient conjugué par rapport à Gauss-Seidel

On peut constater que la méthode du gradient conjugué reste beaucoup plus efficace, même pour de très grandes matrices. Techniquement, le gradient conjugué est un exemple de méthode d'espace de Krylov. Un élément crucial des deux méthodes est la minimisation d'une mesure d'erreur à la k -ième itération. Une autre méthode de ce type est la méthode GMRES (méthode du résidu minimal généralisé). Cette méthode GMRES peut être utilisée pour des systèmes non symétriques, mais il faut stocker une base pour le k -ième espace de Krylov K_k et l'algorithme doit être redémarré à un moment donné, ce qui augmente le temps de convergence. La k -ième itération de GMRES est la solution d'un problème de moindres carrés :

$$\text{minimiser}_{x \in x_0 + K_k} \|b - Ax\|_2.$$

Les méthodes itératives sont généralement efficaces à condition que le problème soit préconditionné (le préconditionneur transformant le problème en un problème plus adapté aux méthodes itératives). En pratique, le préconditionnement consiste souvent à tenter de réduire le nombre de conditionnement, qui mesure la sensibilité d'une fonction aux variations de la donnée en entrée. Réduire le nombre de conditionnement améliore les performances de l'itération, tout comme transformer le problème de sorte que les valeurs propres soient regroupées près de 1.

Par exemple, lors d'une itération, nous pouvons souhaiter remplacer le système linéaire

$$Ax = b$$

dans la méthode du gradient conjugué par un autre système symétrique défini positif et ayant la même solution. Nous pourrions le faire en exprimant le problème préconditionné en termes d'une nouvelle matrice symétrique définie positive B telle que

$$A = B^2$$

et en utilisant ensuite un préconditionneur bilatéral qui approxime B^{-1} . Le carré de ce préconditionneur est égal à M , où M est une matrice symétrique définie positive proche de A^{-1} et nous pouvons donc maintenant exprimer le système préconditionné en termes de la matrice SAS dont

3. Jamie Graves, "Techniques de calcul pour les matrices creuses", Actes de la conférence Tomorrow's Mathematicians Today, Manchester Metropolitan University (2017).

les valeurs propres sont regroupées autour de 1^4 .

Il est également possible d'étudier de grands systèmes denses, par opposition à de grands systèmes creux. Les méthodes de résolution de ces systèmes sont évidemment différentes, puisque l'opération principale dans le cas creux est la multiplication matrice-vecteur, qui n'aura qu'un coût de calcul linéaire pour un système creux. Une façon de traiter un système dense consiste à remplacer la matrice originale A par une matrice approchée \tilde{A} telle que le coût de calcul d'un produit matrice-vecteur soit bien inférieur à n^2 , c'est-à-dire que le coût de calcul soit raisonnable. Plus précisément, le coût de calcul d'une multiplication matrice-vecteur avec la nouvelle matrice approchée doit être bornée par $O(n \log^\alpha n)$, la mémoire nécessaire pour stocker la matrice et le temps requis pour générer la matrice doivent tous deux être bornés par $O(n \log^\beta n)$, et pour un certain n positif et une norme matricielle $\|\cdot\|$, la matrice approchée doit satisfaire

$$\|A - \tilde{A}\| \leq \varepsilon.$$

Dans les deux cas, α et β ne doivent pas dépendre de n .

Vous pourriez vous demander quel type de problème produirait des matrices denses. Lorsque nous interpolons avec des fonctions de base radiales, nous formons un interpolant de la forme

$$s(x) = \sum_{j=1}^n \alpha_j K(x, x_j),$$

où le noyau est défini positif

$$K(x, y) = \Phi(x - y) = \phi(\|x - y\|_2).$$

Les conditions d'interpolation créent un système

$$A\alpha = f$$

avec une matrice d'interpolation A qui est symétrique définie positive. Cette matrice sera positive lorsque le noyau n'a pas de support local. Dans ce cas, il se trouve que le coût de calcul des entrées de la matrice est faible malgré la densité de la matrice, mais ce n'est pas un exemple typique⁵.

Un autre classique de l'algèbre linéaire numérique que nous pourrions mentionner est l'élimination de Gauss. Cette méthode était connue des algébristes linéaires chinois et redécouverte par Gauss. Le résultat était également connu de Newton (comme toute méthode d'algèbre élémentaire ou d'algèbre linéaire). L'élimination de Gauss fournit un algorithme pour résoudre des systèmes d'équations linéaires simultanées. La procédure générale est la suivante : prenez un système de n équations linéaires avec une matrice de coefficients A . Écrivez la matrice augmentée $A|b$ avec n lignes R_n . Soustrayez des multiples de la première ligne des autres lignes jusqu'à ce que les éléments sous la diagonale principale soient nuls dans la première colonne. Répétez le processus avec

4. C. Kelly, Méthodes itératives pour les équations linéaires et non linéaires (Philadelphie : Société de mathématiques industrielles et appliquées, 1995).

5. Holger Wendland, Numerical Linear Algebra : An Introduction (Cambridge : Cambridge University Press, 2017).

des multiples de la deuxième ligne jusqu'à ce que les éléments sous la diagonale principale soient nuls dans la deuxième colonne.

Continuez ainsi jusqu'à obtenir une matrice augmentée telle que la matrice des coefficients soit triangulaire supérieure. Résolvez le système d'équations correspondant en remplaçant toutes les valeurs à l'envers. Nous allons montrer à quoi cela pourrait ressembler dans un exemple explicite. Commencez par un système d'équations linéaires :

$$\begin{aligned}x - 2y + 5z &= 6, \\x + 3y - 4z &= 7, \\2 + 6y - 12z &= 12.\end{aligned}$$

Nous devons faire disparaître les éléments situés sous la diagonale principale de la première colonne. Cela peut être fait en effectuant des opérations élémentaires sur la première ligne : par exemple, soustraire la première ligne de la deuxième ligne serait la chose évidente à faire. Pour la troisième ligne, vous devez multiplier la première ligne par 2 avant de la soustraire. Nous avons alors

$$\begin{aligned}x - 2y + 5z &= 6, \\5y - 9z &= 1, \\10y - 22z &= 0.\end{aligned}$$

Je ne vais pas me donner la peine d'écrire tout cela pour que les coefficients soient des entrées d'une matrice augmentée, mais vous pouvez le faire si vous le trouvez utile. Répétez le processus en effectuant des opérations élémentaires avec la deuxième ligne jusqu'à ce que l'élément situé sous la diagonale principale de la deuxième colonne s'annule.

$$\begin{aligned}x - 2y + 5z &= 6, \\5y - 9z &= 1, \\-4z &= -2.\end{aligned}$$

Les lignes de A sont linéairement indépendantes, il existe donc une solution unique à ce système, qui peut être trouvée trivialement. Cette méthode peut échouer de diverses manières et ne fonctionnera pas toujours, mais je n'entrerai pas dans les détails, car on peut la trouver dans n'importe quel livre d'algèbre linéaire. Une application de la méthode est l'interpolation polynomiale.

Dans le domaine de l'algèbre linéaire, Olga Holtz a étudié les conditions sous lesquelles une matrice de fonctions différentiables doit commuter avec sa dérivée élément par élément. Elle a posé le problème de savoir si une matrice sur un corps différentiel \mathbb{F} avec un corps de constantes algébriquement clos tel que

$$MM' = M'M$$

doit avoir des valeurs propres qui sont des éléments du corps \mathbb{F} . Un corps différentiel est un corps algébrique muni d'une opération supplémentaire appelée dérivée qui satisfait

$$\begin{aligned}(a + b)' &= a' + b', \\(ab)' &= ab' + a'b.\end{aligned}$$

Olga Holtz démontre également divers autres résultats relatifs à la triangularisabilité et à la diagonalisabilité des matrices dans ce contexte : par exemple, si \mathbb{F} est un corps différentiel avec un corps

de constantes algébriquement clos \mathbb{K} et si $M \in \mathbb{F}^{n,n}$ est de type 1, alors M est \mathbb{K} -triangularisable.

Ce théorème implique que les matrices de type 1 ont n valeurs propres dans \mathbb{F} si \mathbb{K} est algébriquement clos. Une matrice de type 1 est une matrice de la forme

$$M_1 = \sum_{\lambda} f_{\lambda} C_{\lambda},$$

où les C_{λ} sont des matrices constantes commutant deux à deux. Une matrice de type 2 est une matrice de la forme

$$M_2 = (f_{\alpha} g_{\beta}),$$

où $f_1, \dots, f_n, g_1, \dots, g_n$ sont des fonctions arbitraires satisfaisant les conditions

$$\sum_{\alpha} f_{\alpha} g_{\alpha} = \sum_{\alpha} f'_{\alpha} g_{\alpha} = 0,$$

ce qui implique que

$$\sum_{\alpha} f_{\alpha} g'_{\alpha} = 0.$$

Cette distinction se trouve dans une lettre de Schur⁶.

Étroitement liée à l'analyse numérique se trouve une vaste théorie appelée théorie de l'approximation. Considérons une courbe du plan donnée par une fonction, puis réfléchissons à la manière de tracer une droite qui s'ajuste au mieux à cette courbe. Il s'agit du problème prototypique de la théorie de l'approximation. On dispose d'une fonction ou d'un élément d'un ensemble à approximer, d'un ensemble d'approximations possibles \mathcal{A} (l'ensemble de toutes les droites dans l'exemple donné), et d'une méthode pour sélectionner la meilleure approximation parmi l'ensemble \mathcal{A} . Comment mesurer la qualité d'une approximation? Rappelons que les espaces métriques sont munis d'une fonction de distance qui peut être utilisée à cette fin, et il arrive souvent, lorsqu'on travaille sur un problème d'approximation, qu'il existe un espace métrique approprié contenant à la fois la fonction à approximer et l'ensemble des approximations \mathcal{A} . On pourrait dire que a_1 est une meilleure approximation que a_2 si l'inégalité

$$d(a_1, f) < d(a_2, f)$$

est satisfaite. On peut définir a_B comme étant la meilleure approximation si l'inégalité

$$d(a_B, f) < d(a, f)$$

est valable pour tout $a \in \mathcal{A}$. N'oubliez pas que l'espace métrique doit pouvoir mesurer l'erreur lorsque nous tentons une approximation d'essai afin d'améliorer nos tentatives pour obtenir un meilleur ajustement. Certains résultats célèbres que vous connaissez peut-être (le théorème de Stone-Weierstrass et la formule d'interpolation de Lagrange, par exemple) sont des résultats de la théorie de l'approximation⁷.

6. Olga Holtz, Volker Mehrmann et Hans Schneider, "Matrices qui commutent avec leur dérivée. Sur une lettre de Schur à Wielandt", *Linear Algebra Appl.* 438(5), 2574-2590 (2012).

7. Michael Powell, *Approximation Theory and Methods* (Cambridge : Cambridge University Press, 1981).