1 Fonctions de matrices

1.6 Meilleure approximation

Soit $\mathbb{E} \subset \mathbb{C}$ un convexe compact et f analytique dans un voisinage de $\mathbb{E}$. Dans ce chapitre on cherche à minimiser $\|f-p\|_{\mathbb{E}}$ pour un polynôme de degré $\leq n$ (ou une fonction rationnelle à pôles fixes). On note par $\phi : \mathbb{C} \setminus \mathbb{E} \mapsto \mathbb{C} \setminus \mathbb{D}$ l'application de Riemann (l'unique bijection analytique conforme vérifiant $\phi(\infty) = \infty, \phi'(\infty) > 0$ et $\forall\, z : \phi'(z) \neq 0$), et $\psi = \phi^{-1}$. L'ensemble de niveau $\mathbb{E}_R$ pour $R > 1$ est défini par son complément $\mathbb{E}_R^c = \{z \notin \mathbb{E} : |\phi(z)| > R\}$.

1.6.1 Définition : On définit $F_j(z)$ pour $z \in \mathrm{int}(\mathbb{E})$, $|w| \geq 1$ (ou $z \in \mathbb{E}$, $|w| > 1$) par la fonction génératrice

$$\frac{w\psi'(w)}{\psi(w) - z} = \sum_{j=0}^{\infty} \frac{F_j(z)}{w^j}.$$

Pour l'exemple $\mathbb{E} = \mathbb{D}$, nous avons $\psi(w) = w$, et $F_j(z) = z^j$. Pour l'exemple $\mathbb{E} = [-1, 1]$, $\psi(w) = \frac{1}{2}(w + \frac{1}{w})$, et $F_0(z) = 1$ et pour $j \geq 1 : F_j(\psi(w)) = w^j + \frac{1}{w^j} = 2T_j(\psi(w))$.

1.6.2 Lemme : F_j est un polynôme de degré j, $F_0(z) = 1$, et pour $j \geq 1 : F_j(\psi(w)) - w^j$ est analytique dans $|w| > 1$ inclus ∞ et s'annule en ∞.
Preuve : La série génératrice étant absolument convergente pour $|w| = 1 + \epsilon > 1$, on obtient pour $k \in \mathbb{Z}$, $z \in \mathbb{E}$

$$\begin{aligned}(*) \qquad & \frac{1}{2\pi i} \int_{|w|=1} w^k \frac{w\psi'(w)}{\psi(w) - z} \frac{dw}{w} \\ &= \sum_{j=0}^{\infty} F_j(z) \frac{1}{2\pi i} \int_{|w|=1+\epsilon} w^{k-j} \frac{dw}{w} = \begin{cases} F_k(z) & k \geq 0, \\ 0 & k < 0. \end{cases}\end{aligned}$$

en particulier en écrivant $\phi(\zeta)^j - P(\zeta)$ analytique en $\mathbb{C}\setminus\mathbb{E}$ et s'annulant en ∞ avec P un polynôme de degré j

$$F_j(z) - P(z) = \frac{1}{2\pi i} \int_{\partial\mathbb{E}} (\phi(\zeta)^j - P(\zeta)) \frac{d\zeta}{\zeta - z} = 0$$

d'après le théorème de Cauchy.

Voici un résultat utilisant la convexité de $\mathbb{E}$.
1.6.3 Définition et Théorème : Pour P polynôme et $z \in \mathrm{int}(\mathbb{E})$, soit

$$\mathcal{F}(P)(z) = \frac{1}{2\pi i} \int_{|w|=1} P(w)\, 2\mathrm{Re}\Big(\frac{w\psi'(w)}{\psi(w) - z}\Big) \frac{dw}{w}.$$

(a) $\mathcal{F}(1)(z) = 2$, et pour $j \geq 1 : \mathcal{F}(w^j)(z) = F_j(z)$.
(b) $\|\mathcal{F}(P)\|_{\mathbb{E}} \leq 2\, \|P\|_{\mathbb{D}}$, en particulier $\|F_j\|_{\mathbb{E}} \leq 2$.
(c) $\mathcal{F}(P)(\psi(w)) - P(w)$ est analytique dans $|w| > 1$ inclus ∞.
Preuve : Pour $j \geq 0$

$$\mathcal{F}(w^j)(z) = \frac{1}{2\pi} \int_{|w|=1} w^j \frac{w\psi'(w)}{\psi(w) - z} \frac{dw}{iw} + \overline{\frac{1}{2\pi} \int_{|w|=1} w^{-j} \frac{w\psi'(w)}{\psi(w) - z} \frac{dw}{iw}}$$

ce qui d'après $(*)$ vaut 2 si $j = 0$, et $F_j(z)$ pour $j > 0$, ce qui démontre (a). Pour une preuve de la partie (b), notons d'abord que pour $w = e^{it}$ nous avons $\frac{dw}{iw} = dt > 0$. Aussi, on montre que

$w\psi'(w)/|w\psi'(w)|$ nous donne la normale extérieure de $\mathbb{E}$ au point $z = \psi(w)$. Donc par convexité pour $z \in \mathrm{int}(\mathbb{E})$

$$\mathrm{Re}\Big(\frac{w\psi'(w)}{\psi(w) - z}\Big) > 0,$$

ce qui permet d'estimer

$$|\mathcal{F}(P)(z)| \leq \frac{\|P\|_{\mathbb{D}}}{2\pi} \int_{|w|=1} \Big| 2\mathrm{Re}\Big(\frac{w\psi'(w)}{\psi(w) - z}\Big) \frac{dw}{iw} \Big| = \|P\|_{\mathbb{D}} \mathcal{F}(1) = 2\,\|P\|_{\mathbb{D}}.$$

1.6.4 Exercice : En suivant le raisonnement de la preuve du théorème de Neumann, montrer que $W(A) \subset \mathbb{E}$ et $p = \mathcal{F}(P)$ pour un polynôme P implique que $\|p(A)\| \leq 2\,\|P\|_{\mathbb{D}}$, et en particulier $\|F_j(A)\| \leq 2$.

1.6.5 Exercice : Soit f analytique dans un voisinage de $\mathbb{E}_R$ pour $R > 1$, alors avec

$$f_j := \frac{1}{2\pi i} \int_{|w|=1} \frac{f(\psi(w))}{w^j} \frac{dw}{w}$$

dits coefficients de Faber montrer que $f_j = \mathcal{O}(R^{-j})_{j\to\infty}$, et que les sommes partielles de la somme de Faber $\sum_{j=0}^{\infty} f_j F_j(z)$ convergent vers f uniformément dans $\mathbb{E}$.

L'exo 1.6.5 nous permet d'étendre la définition de $\mathcal{F}$ à tout P analytique dans un voisinage de $\mathbb{D}$, tout en gardant les propriétés 1.6.3(b),(c), et

$$\mathcal{F}\Big(\frac{f_0}{2} + \sum_{j=1}^{\infty} f_j w^j\Big)(z) = \sum_{j=0}^{\infty} f_j F_j(z).$$

On a le résultat suivant concernant la meilleure approximation polynômiale.

1.6.6 Corollaire : Soit f analytique dans un voisinage de $\mathbb{E}$, alors

$$|f_{m+1}| \leq \sqrt{\sum_{j=m+1}^{\infty} |f_j|^2} \leq \min_{P \in \mathcal{P}_m} \|f - P\|_{\mathbb{E}} \leq \|f - \sum_{j=0}^{m} f_j F_j\|_{\mathbb{E}} \leq 2 \sum_{j=m+1}^{\infty} |f_j|.$$

Ce corollaire 1.6.6 nous donne un encadrement précis et un "bon" approximant explicite si les f_j décroissent rapidement, voir l'exemple suivant. Notons que la première et troisième inégalité sont évidentes, et la quatrième découle de l'estimation de $\|F_j\|_{\mathbb{E}}$ donnée dans 1.6.3(b). Une preuve de la deuxième inégalité va nous demander un peu d'effort, elle découlera comme cas particulier du théorème 1.6.8 ci-dessous.

1.6.7 Corollaire : Soit $\mathbb{E}$ symétrique par rapport à l'axe réelle, et $[a, b] \subset \mathbb{R}$, à gauche de $\mathbb{E}$ (c'est-à-dire, $b < \min\{\mathrm{Re}(z) : z \in \mathbb{E}\}$), alors pour la fonction de Markov

$$f(z) = \int_a^b \frac{d\mu(x)}{z - x},$$

pour $j \geq m + 1$ nous avons

$$|f_j| \leq |\phi(b)|^{m+1-j}|\,|f_{m+1}|, \qquad \sum_{j=m+1}^{\infty} |f_j| \leq \frac{2}{|\phi(b)|^{m+1}}\,\|f\|_{\mathbb{E}}$$

($\Longrightarrow$ l'estimation du 1.6.6 est précise à un facteur $2/(1-|\phi(b)|^{-1})$ près).
Preuve : D'après le théorème de Fubini nous avons

$$f_j \;=\; \frac{1}{2\pi i}\int_{|w|=1}\int_a^b \frac{d\mu(x)}{\psi(w)-x}\frac{dw}{w^{j+1}} = \int_a^b\int_a^b d\mu(x)\frac{1}{2\pi i}\int_{|w|=1}\frac{1}{\psi(w)-x}\frac{dw}{w^{j+1}}.$$

L'intégrant de l'intégrale en w admet une seule singularité dans $\mathbb{D}^c \cup \{\infty\}$, au point $w=\phi(x)$. Donc, par le théorème des résidus en analyse complexe (ou tout simplement par le théorème de Cauchy après changement de variables $\zeta=\psi(w)$ et changement d'orientation de la courbe d'intégration),

$$\frac{1}{2\pi i}\int_{|w|=1}\frac{1}{\psi(w)-x}\frac{dw}{w^{j+1}} = -\frac{1}{\psi'(\phi(x))}\frac{1}{\phi(x)^{j+1}} = -\frac{\phi'(x)}{\phi(x)^{j+1}}.$$

Par unicité de l'application de Riemann et symétrie de $\mathbb{E}$, nous avons $\phi(\overline{z})=\overline{\phi(z)}$ pour tout $z\notin\mathbb{E}$, en particulier, $\phi(x)$ et $\phi'(x)\neq 0$ sont réels pour $x\in\mathbb{R}\setminus\mathbb{E}\supset[a,b]$. Comme de plus $\phi'(\infty)>0$, $\phi(\infty)=\infty$, nous déduisons que $\phi'>0$ dans $\mathbb{R}\setminus[a,b]$, et donc ϕ est croissant et négatif sur $[a,b]$, et $1/|\phi|$ croit sur $[a,b]$. Donc

$$|f_j| \;=\; \Big|\int_a^b \frac{\phi'(x)}{\phi(x)^{j+1}}d\mu(x)\Big| = \int_a^b \frac{|\phi'(x)|}{|\phi(x)|^{j+1}}d\mu(x) \leq |\phi(b)|^{m+1-j}|\,|f_{m+1}|.$$

Du théorème 3.1 de l'article [K. C. Toh and L. N. Trefethen, The Kreiss matrix theorem on a general complex domain, SIAM J. Matrix Anal. Appl., 21 (1999), pp. 145–165] on sait que, pour tout domaine $\mathbb{E}$ simplement connexe pas forcement convexe,

$$\forall z\notin\mathbb{E}:\quad \mathrm{dist}(z,\mathbb{E})\frac{|\phi'(z)|}{|\phi(z)|-1}\in[\frac{1}{2},2].$$

Par conséquent,

$$\sum_{j=m+1}^{\infty}|f_j| = \int_a^b \frac{|\phi'(x)|\,d\mu(x)}{(1-|\phi(x)|^{-1})|\phi(x)|^{m+2}} \leq \frac{2}{|\phi(b)|^{m+1}}\int_a^b\frac{d\mu(x)}{\mathrm{dist}(x,\mathbb{E})} = \frac{2}{|\phi(b)|^{m+1}}\,\|f\|_{\mathbb{E}}.$$

Nous allons maintenant démontrer un résultat similaire à 1.6.6 pour les fonctions rationnelles à pôles prescrits. Pour $w_1,...,w_m\in\mathbb{C}\setminus\mathbb{D}$ soit $Q(w)=\prod_{j=1}^m(1-w/w_j)$, et $q(z)=\prod_{j=1}^m(z-\psi(w_j))$. On va supposer dans la suite que les w_j (et donc les $z_j=\psi(w)$) soient distincts. Néanmoins, les idées de preuve restent valables après des passages à la limite, par exemple $w_1\to w_2$, mais aussi $w_1\to\infty$ (et donc $1-w/w_1\to 1$, ce qui veut dire que Q sera de degré $<m$). En particulier, on aura la situation du théorème 1.6.6 en faisant tendre tous les w_j vers ∞.
Rappelons quelques petites éléments de la théorie des espaces de Hardy : on note

$$\|F\|_2 := \sqrt{\frac{1}{2\pi}\int_{|w|=1}|F(w)|^2\,|dw|}$$

pour une fonction F de carré intégrable sur le cercle d'unité. L'identité $\frac{1}{2\pi}\int_{|w|=1}w^{j-k}|dw|=\delta_{j,k}$ plus la théorie des espaces H^2 montre que

$$\|F\|_2 = \sqrt{\sum_{j=-\infty}^{+\infty}|F_j|^2}$$

si F est analytique dans la couronne $1 < |w| < 1+\epsilon$ et y admet alors un développement de Laurent $F(w) = \sum_{j=-\infty}^{+\infty} F_j w^j$.

1.6.8 Théorème : Soit f analytique dans un voisinage de $\mathbb{E}$. Notons $R_m = P_m/Q$ l'interpolant de $F(w) = f_0/2 + \sum_{j=1}^{\infty} f_j w^j$ aux points 0 et $1/\overline{w}_1, ..., 1/\overline{w}_m$, et

$$B(w) := w \prod_{j=1}^{m} \frac{w - 1/\overline{w}_j}{1 - w/w_j}, \quad \frac{p_m}{q} := \mathcal{F}(\frac{P_m}{Q}), \quad b_j := \frac{1}{2\pi i} \int_{|u|=1} \frac{f(\psi(w))}{B(u)} \frac{du}{u^j}.$$

Alors

$$|b_1| \leq \sqrt{\sum_{j=1}^{\infty} |b_j|^2} \leq \min_{p \in \mathcal{P}_m} \|f - \frac{p}{q}\|_{\mathbb{E}} \leq \|f - \frac{p_m}{q}\|_{\mathbb{E}} \leq 2 \sum_{j=1}^{\infty} |b_j|.$$

Avant de se lancer dans la preuve, notons que pour $w_1, ..., w_m \to \infty$, B devient w^{m+1} et donc $b_j = f_{j+m}$. Aussi, $P_m/Q = P_m$ devient la somme partielle de F d'ordre m, donc le corollaire 1.6.6 est en effet un cas limite du théorème 1.6.8.

Preuve : Dans un premier temps, montrons que $p_m \in \mathcal{P}_m$, c'est-à-dire, p_m/q est effectivement un candidat pour notre problème de minimisation. En écrivant la décomposition en termes simples et en utilisant la fonction génératrice des polynômes de Faber nous obtenons

$$\begin{aligned}
\mathcal{F}\Big(\frac{P_m(w)}{Q(w)}\Big)(z) &= \mathcal{F}\Big(c_0 + \sum_{j=1}^{m} \frac{c_j}{w - w_j}\Big)(z) = \mathcal{F}\Big(c_0 - \sum_{j=1}^{m} \frac{c_j}{w_j} \sum_{k=0}^{\infty} \frac{w^k}{w_j^k}\Big)(z) \\
&= c_0 \mathcal{F}(1)(z) - \sum_{j=1}^{m} \frac{c_j}{w_j} \sum_{k=0}^{\infty} \frac{\mathcal{F}(w^k)(z)}{w_j^k} \\
&= c_0 + \frac{P_m}{Q}(0) - \sum_{j=1}^{m} \frac{c_j}{w_j} \frac{w_j \psi'(w_j)}{\psi(w_j) - z} = c_0 + \frac{P_m}{Q}(0) - \sum_{j=1}^{m} \frac{c_j \psi'(w_j)}{z - \psi(w_j)}
\end{aligned}$$

étant clairement un élément de $\mathcal{P}_m/q$. D'ailleurs, cette formule très explicite permet de construire sur ordinateur p_m/q sachant la décomposition en termes simples de P_m/Q.
On passe maintenant à une preuve de la troisième inégalité sachant que la deuxième est triviale. Observons d'abord que $F - P_m/Q$ est analytique dans un voisinage de $|w| \leq 1+\epsilon$ pour un $\epsilon > 0$. D'après la formule d'Hermite 1.5.4, nous obtenons pour $|w| = 1$ sachant que $|B(w)| = 1$

$$\begin{aligned}
\Big|F(w) - \frac{P_m}{Q}(w)\Big| &= |B(w)| \Big|\frac{1}{2\pi i} \int_{|u|=1+\epsilon} \frac{F(u)}{B(u)} \frac{du}{u - w}\Big| \\
&= \Big|\frac{1}{2\pi i} \int_{|u|=1+\epsilon} \frac{f(\psi(u))}{B(u)} \frac{du}{u - w}\Big| = \Big|\sum_{j=1}^{\infty} b_j w^{j-1}\Big| \leq \sum_{j=1}^{\infty} |b_j|,
\end{aligned}$$

où dans la deuxième égalité on a utilisé le fait que

$$u \mapsto \frac{F(u) - f(\psi(u))}{B(u)} \frac{1}{u - w}$$

est analytique dans $|u| \geq 1+\epsilon$ inclus ∞, voire 1.6.3(c), avec un double zéro en ∞. En utilisant 1.6.3(b), on en déduit que

$$\|f - \frac{p_m}{q}\|_{\mathbb{E}} = \|\mathcal{F}(F - \frac{P_m}{Q})\|_{\mathbb{E}} \leq 2\, \|F - \frac{P_m}{Q}\|_{\mathbb{D}} \leq 2 \sum_{j=1}^{\infty} |b_j|,$$

c'est-à-dire, il reste seulement la première inégalité à établir.
Pour tout $p \in \mathcal{P}_m$ nous pouvons écrire

$$(f - \frac{p}{q})(\psi(w)) = w(\widetilde{F}(w) - \frac{P}{Q}(w)) + H(w), \quad \widetilde{F}(w) = \sum_{j=1}^{\infty} F_j w^j = \frac{F(w) - F(0)}{w},$$

avec $P \in \mathcal{P}_{m-1}$, et H analytique dans $|u| > 1$ d'après 1.6.3(c) (développer $f - p/q$ en série de Faber). Comme le terme à gauche du second membre est analytique dans un voisinage du disque, et s'annule en 0, nous obtenons alors

$$\|f - \frac{p}{q}\|_{\mathbb{E}}^2 = \|(f - \frac{p}{q}) \circ \psi\|_{\partial\mathbb{D}}^2 \geq \|(f - \frac{p}{q}) \circ \psi\|_2^2 = \|\widetilde{F} - \frac{P}{Q}\|_2^2 + \|H\|_2^2.$$

Notons qu'il existe un polynôme $\widetilde{P} \in \mathcal{P}_{m-1}$ de sorte que

$$\frac{P_m}{Q} - F(0) = \frac{P_m}{Q} - \frac{P_m}{Q}(0) = w\,\frac{\widetilde{P}}{Q} \quad \text{et alors} \quad \|\widetilde{F} - \frac{\widetilde{P}}{Q}\|_2^2 = \|F - \frac{P_m}{Q}\|_2^2 = \sum_{j=1}^{\infty} |b_j|^2,$$

la dernière égalité découlant de la représentation intégrale de $|F - P/Q|$ donnée ci-dessus. En effet, le lecteur vérifie aisément que $\widetilde{P}/Q$ n'est rien que l'interpolant $\in \mathcal{P}_{m-1}/Q$ de $\widetilde{F}$ aux points $1/\overline{w}_1, ..., 1/\overline{w}_m$. En combinant ces deux chaînes d'inégalités, il est suffisant de démontrer que

$$\|\widetilde{F} - \frac{\widetilde{P}}{Q}\|_2 = \min_{P \in \mathcal{P}_{m-1}} \|\widetilde{F} - \frac{P}{Q}\|_2,$$

autrement dit, on connaît le meilleur approximant par rapport à la norme $\|\cdot\|_2$ induite par un produit scalaire

$$\langle G, H\rangle = \frac{1}{2\pi}\int_{|w|=1} F(w)\overline{G(w)}|dw| = \frac{1}{2\pi i}\int_{|w|=1} F(w)\overline{G(w)}\frac{dw}{w},$$

défini, disons, sur l'espace vectoriel des fonctions analytiques dans un voisinage fixe de $\mathbb{D}$ (c'est en effet le produit scalaire de l'espace plus grand H^2 de Hardy). Ceux qui étaient en M315 en L3 avec moi savent qu'il est bien plus "sympa" de minimiser au sens des moindres carrés :

$$\frac{\widetilde{P}}{Q}(w) = \sum_{j=1}^{m} \frac{e_j}{w - w_j}$$

est meilleur approximant par rapport à $\|\cdot\|_2$ de $\widetilde{F}$ si et seulement si l'erreur $\widetilde{F} - \frac{\widetilde{P}}{Q}$ est orthogonal à toute fonction dans $\mathcal{P}_{m-1}/Q$, avec base $1/(w - w_\ell)$, $\ell = 1, ..., m$. Il faut et il suffit alors que

$$\Big[\ \langle \widetilde{F}, \tfrac{1}{w-w_\ell}\rangle\ \Big]_{\ell=1,\dots,m} = \Big[\ \langle \tfrac{1}{w-w_j}, \tfrac{1}{w-w_\ell}\rangle\ \Big]_{\ell,j=1,\dots,m} \big[\ c_j\ \big]_{j=1,\dots,m}.$$

Un petit calcul de résidus montre que

$$\langle \widetilde{F}, \frac{1}{w - w_\ell}\rangle = -F(1/\overline{w}_\ell)/\overline{w}_\ell, \qquad \langle \frac{1}{w - w_j}, \frac{1}{w - w_\ell}\rangle = -\frac{1}{1/\overline{w}_\ell - w_j}/\overline{w}_\ell$$

et donc notre système est équivalent au fait que $\widetilde{P}/Q$ interpole $\widetilde{F}$ aux points $1/\overline{w}_1, ..., 1/\overline{w}_m$, comme désiré ci-dessus.

1.6.9 Exercice : Avec les notations de 1.6.8, si $W(A) \subset \mathbb{E}$ alors

$$\|f(A) - p_m(A)q(A)^{-1}\| \leq \|F - P_m/Q\|_{\mathbb{D}} \leq \sum_{j=1}^{\infty} |b_j|.$$

Références

Pour les espaces de Hardy : E. M. Nikishin, V. N. Sorokin, Rational Approximations and Orthogonality, *Transl. Amer. Math. Soc.*, Vol. **92**, Providence, R.I. (1991). (Bibmath)

Pour les polynômes de Faber (attention, la normalisation des polynômes de Faber est un peu différente) : P. Henrici, Applied and Computational Complex Analysis, vol. 3, Wiley, New York, 1974. (Bibmath et books.google.fr) Chapitre 18.

Et pour le sujet du chapitre (à télécharger sur mon site web) : B. Beckermann, L. Reichel : Error estimation and evaluation of matrix functions via the Faber transform, SIAM J. Num. Anal. 47 (2009), 3849-3883.

B. Beckermann, Image numérique, GMRES et polynômes de Faber, C. R. Acad. Sci. Paris, Ser. I 340(2005) 855-860.