§ 1.4. Estimations pour fonctions de matrices

Pour pouvoir approcher $f(A)$ par $g(A)$ pour une fonction g (rationnelle ou polynomiale) "proche" de f, la notion suivante est utile:

<u>1.4.1. Déf</u>: Un ensemble fermé $\Omega \subset \mathbb{C}$ est dit K-spectral pour $A \in \mathbb{C}^{n\times n}$ si pour tout f analytique dans un voisinage de Ω : $\|f(A)\| \leq K \cdot \underbrace{\sup_{z\in\Omega} |f(z)|}_{=:\|f\|_\Omega}$. □

Notons que forcément $\sigma(A) \subset \Omega$ (on peut sinon construire une suite de fonctions f_n restant bornée sur Ω et en tout point de $\sigma(A)$ sauf un).

<u>1.4.2. Lemme</u>: $\sigma(A)$ est K-spectral avec $K = \mathrm{cond}(Z) = \|Z\|\cdot\|Z^{-1}\|$ si A est diagonalisable: $Z^{-1}AZ \overset{D}{=} \mathrm{diag}$.

<u>Preuve</u>: $\|f(A)\| = \|Z\cdot f(D)\cdot Z^{-1}\| \leq \mathrm{cond}(Z)\cdot \underbrace{\|\mathrm{diag}(f(\lambda_j))\|}_{=\|f\|_{\sigma(A)}}$ ∎

L'hypothèse diagonalisable est essentielle: voir $f(z) = \lambda z$, $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $\|f(A)\| = |\lambda|$, $\|f\|_{\sigma(A)} = 0$, on pourrait seulement donner une borne en fonction des dérivées de f ... sinon il faut choisir Ω plus grand!

Notons que si ϕ est analytique dans un voisinage de Ω alors $\phi(\Omega)$ est aussi K-spectral pour $\phi(A)$.

<u>1.4.3. Théorème de Neumann</u>

Le disque $\Omega = \{z \in \mathbb{C} : |z-\omega| \leq R\}$ avec $R \geq \|A-\omega I\|$ est 1-spectral pour A (voir exemple: $\omega=0$, $R=1$, $\Omega=\mathbb{D}$).

<u>Preuve</u> (voir Riesz-Nagy, §154) On supposera d'abord que $R > \|A-\omega I\|$, et ensuite on fera tendre R vers $\|A-\omega I\|$.
En posant $U = \frac{A-\omega I}{R}$, $g(z) = f(\omega + R\cdot z)$ nous avons $\|U\| < 1$, g analytique dans voisinage de $\mathbb{D}$ (si R assez petit)

et alors d'après 1.1.11

$$f(A) = g(U) = \frac{1}{2\pi i} \int_{\partial \mathbb{D}} g(w)\,(wI-U)^{-1} \underbrace{dw}_{= i\cdot w\,|dw|}$$

$$= \frac{1}{2\pi} \int g(w) \left[\tfrac{1}{2} I + \tfrac{1}{2}(wI+U)(wI-U)^{-1} \right] |dw|$$

$$= \frac{g(0)}{2} I + \frac{1}{2\pi} \int g(w) \left[\tfrac{1}{2}(wI+U)(wI-U)^{-1} \right] |dw|$$

aussi $\frac{1}{2\pi} \int_{|w|=1} g(w) \left[\frac{1}{2}(wI+U)(wI-U)^{-1} \right]^{*} |dw|$

$$= \frac{1}{2\pi i} \int g(w) \left[\tfrac{1}{2} (I - wU^{*})^{-1}(I + wU^{*}) \right] \frac{dw}{w} = \frac{g(0)}{2} I$$

d'après le théorème des résidus (ou série de Neumann).

Donc $f(A) = \frac{1}{2\pi} \int_{|w|=1} g(w)\, M(w)\, |dw|$

avec $M(w) = \mathrm{Re}\left(\left(I + \frac{U}{w}\right)\left(I - \frac{U}{w}\right)^{-1} \right)$ hermitienne déf. positive, car avec $y = \left(I - \frac{U}{w}\right)\tilde{y} \in \mathbb{C}^n$

$$y^{*} M(w)\, y = \mathrm{Re}\ \tilde{y}^{*} \left(I - \frac{U}{w}\right)^{*} \left(I + \frac{U}{w}\right) \tilde{y}$$

$$= \tilde{y}^{*}\tilde{y} - \tilde{y}^{*} U^{*} U \tilde{y} \geq 0 \quad \text{car } \|U\| < 1.$$

donc $\forall \|x\|,\|y\| \le 1$: $\max_{x,y} |x^{*} f(A)\, y| \leq \|g\|_{\mathbb{D}} \cdot \frac{1}{2\pi} \int_{|w|=1} \sqrt{x^{*}M(w)x}\, \sqrt{y^{*}M(w)y}\, |dw|$

$$\leq \|g\|_{\mathbb{D}} \cdot \sqrt{\frac{1}{2\pi} \int x^{*} M(w)\, x\, |dw|}\ \sqrt{\frac{1}{2\pi} \int y^{*} M(w)\, y\, |dw|}$$

$$= \|g\|_{\mathbb{D}}\ \|x\| \cdot \|y\| \leq \|g\|_{\mathbb{D}} \Rightarrow \|f(A)\| \leq \|g\|_{\mathbb{D}} = \|f\|_{\Omega}.$$

$\blacksquare$

1.4.4 Corollaire:

Si A est semi-déf. positive alors $\Omega = \{\mathrm{Re}(z) \geq 0\}$ est 1-spectral, en particulier $\|\exp(-A)\| \leq 1$.

Preuve : Soit $\phi(z) = \frac{1-z}{1+z}$ alors ϕ est analytique dans Ω, $\phi(\Omega) = \mathbb{D}$, et comme dans la preuve précédente on montre que $\|\phi(A)\| \leq 1$.

Donc $\|f(A)\| = \|(f\circ\phi^{-1})(\phi(A))\| \overset{1.4.3}{\leq} \|f\circ\phi^{-1}\|_{\mathbb{D}} = \|f\|_{\Omega}$.

$f(z) = \exp(-z)$ n'est pas analytique dans Ω, mais $f_t(z) = \exp\left(-\frac{z}{1+tz}\right)$ l'est pour tout $t > 0$, et $\|f_t(A) - f(A)\| \xrightarrow{t \to 0} 0$

$\Rightarrow \|f(A)\| \leq \limsup_{t\to 0} \|f_t(A)\| \leq \limsup_{t\to 0} \|f_t\|_{\Omega} \leq 1.$ $\blacksquare$

Parlons d'une généralisation récente obtenue par Michel Crouzeix en 2007 :

1.4.5. Déf La image numérique (numerical range, field of values) d'une matrice $A \in \mathbb{C}^{n\times n}$ est donnée par
$$W(A) = \left\{ \frac{y^* A y}{y^* y} : y \in \mathbb{C}^n \setminus \{0\} \right\} = \{ y^* A y : \|y\| = 1 \}$$ ▯

On vérifie vite que $W(A)$ est un compact $\subset \{ |z| \le \|A\| \}$ qui contient $\sigma(A)$. Si A est diagonale (normale) alors $W(A)$ est l'enveloppe convexe des éléments digonaux (de $\sigma(A)$). $W\left(\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\right) = \left\{ \frac{\overline{y_1} y_2}{|y_1|^2 + |y_2|^2} : y_1, y_2 \in \mathbb{C} \right\} = \frac{1}{2}\mathbb{D}$! Voir [Horn et Johnson] !

1.4.6 Lemme : $W(A)$ est convexe.

Preuve : Posons $r_1 = \frac{y_1^* A y_1}{y_1^* y_1}$ et $r_2 = \frac{y_2^* A y_2}{y_2^* y_2}$. Nous devons montrer que $[r_1, r_2] \subset W(A)$, ce qui est trivial si $r_1 = r_2$. Sinon, comme $W(\alpha A + \beta) = \alpha W(A) + \beta$, il suffit de considerer le cas $r_1 = 1$ et $r_2 = -1$ ($A \to \frac{2A - r_1 - r_2}{r_1 - r_2}$). Pour un angle $\varphi \in \mathbb{R}$, soit $y(t) = t \cdot y_1 + (1-t) e^{i\varphi} y_2$. Les matrices $D = \mathrm{diag}(1, e^{i\varphi})$ et $Y = (y_1, y_2)$ sont de rang 2 (sinon $r_1 = r_2$), et alors
$$g(t) = \frac{y(t)^* A y(t)}{y(t)^* y(t)} = \frac{(t, (1-t)) \cdot D^* Y^* A Y D \cdot \binom{t}{1-t}}{(t, (1-t)) \cdot D^* Y^* Y D \binom{t}{1-t}} \qquad g([0,1]) \supset [-1,1]$$
$g(0) = 1$, $g(-1) = -1$, $g : [0,1] \to \mathbb{R}$ continue (pour φ approprié) $\Rightarrow$ ▯

1.4.7. Théorème : $W(A)$ est K-spectral pour A avec $K \le 11.25$ (en général $K \ge 2$, $K = 2$?).

En particulier avec $\Omega = \{ |z| \le \underbrace{\max\{ |\tilde z| : \tilde z \in W(A) \}}_{\text{rayon numérique } \in [\rho(A), \|A\|]} \}$: Ω est 2-spectral pour A.

Retour à notre exemple :
$$\| f(\underbrace{\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}}_{=A,\ W(A) = \frac{1}{2}\mathbb{D}}) \| \le 2 \cdot \max_{|z| \le \frac{1}{2}} |f(z)| \quad (\text{égalité pour } f(z) = \lambda z)$$
$$= \left\| \begin{pmatrix} 0 & f'(0) \\ 0 & 0 \end{pmatrix} \right\| = |f'(0)|$$

Pour toute matrice 2×2, $W(A)$ est une ellipse, mais plus compliqué si $n \ge 3$.

Si A normal, $\|f(A)\| = \|f\|_{\sigma(A)} \leq \|f\|_{\mathrm{conv}(\sigma(A))} = \|f\|_{W(A)}$.

Pour les matrices non normales, on peut aussi considérer

1.4.8. Déf : pseudospectre

Pour $\varepsilon > 0$, soit $\sigma_\varepsilon(A) = \{ z \in \mathbb{C} : \|(zI-A)^{-1}\| \geq \frac{1}{\varepsilon} \}$

(avec convention $\|C^{-1}\| = +\infty$ si C pas inversible

$\Rightarrow \sigma(A) \subset \sigma_\varepsilon(A)$). ▫

Clairement $\sigma_\varepsilon(A)$ croit avec ε. On peut montrer qu'il est compact, mais pas forcément convexe car $\bigcap_{\varepsilon > 0} \sigma_\varepsilon(A) = \sigma(A)$

Voir livre de Trefethen et Embree 2005 Bibmath !

1.4.9 Théorème : Nous avons l'équivalence

(i) $z \in \sigma_\varepsilon(A)$

(ii) $\exists\, y \in \mathbb{C}^n$, $\|y\| = 1$ et $\|(A - zI)\, y\| \leq \varepsilon$

(iii) $\exists\, E \in \mathbb{C}^{n \times n}$, $\|E\| \leq \varepsilon$ avec $z \in \sigma(A+E)$.

Preuve : $\boxed{(i) \Rightarrow (ii)}$:

Soit $v \in \mathbb{C}^n \setminus \{0\}$ avec $\dfrac{\|(zI-A)^{-1} v\|}{\|v\|} = \|(zI-A)^{-1}\| \geq \frac{1}{\varepsilon}$

en posant $y = \dfrac{(zI-A)^{-1} v}{\|(zI-A)^{-1} v\|}$: $\|y\| = 1$ et $\|(zI-A)y\| \leq \varepsilon$.

$\boxed{(ii) \Rightarrow (iii)}$

Posons $(A - zI) \cdot y = \tilde{\varepsilon} \cdot u$ avec $\|u\| = 1$, $|\tilde{\varepsilon}| \leq \varepsilon$.

Alors avec $E = -\tilde{\varepsilon} \cdot u \cdot y^*$: $(A+E)y = Ay - \tilde{\varepsilon} \cdot u = z \cdot y$

$\Rightarrow z \in \sigma(A+E)$, et $\|E\| = \sqrt{\|E^* E\|} = \sqrt{\underbrace{\|y\, y^*\|}_{=1}} \cdot |\tilde{\varepsilon}| \leq \varepsilon$

$\boxed{(iii) \Rightarrow (i)}$ Soit y avec $\|y\| = 1$ et $(A+E)y = zy$.

si $z \notin \sigma(A)$! Donc $y = (zI-A)^{-1} \cdot Ey$

$\Rightarrow \|(zI-A)^{-1}\| \geq \dfrac{\|y\|}{\|Ey\|} \geq \dfrac{\|y\|}{\|E\| \cdot \|y\|} \geq \dfrac{1}{\varepsilon}$. ▫

1.4.10 Exo :

En utilisant que $\|(xI-A)^{-1} - (yI-A)^{-1}\| \leq \|(xI-A)^{-1}\|^2 \dfrac{|x-y|}{1 - |x-y| \cdot \|(xI-A)^{-1}\|}$

pour $x \notin \sigma(A)$, $|x-y| \cdot \|(xI-A)^{-1}\| < 1$, vérifier que

$z \to \|(zI-A)^{-1}\|$ est continue en $z \in \mathbb{C} \setminus \sigma(A)$,

que $\sigma_\varepsilon(A)$ est fermé et que $\partial \sigma_\varepsilon(A) = \{ z \in \mathbb{C} : \|(zI-A)^{-1}\| = \frac{1}{\varepsilon} \}$

1.4.11 Corollaire :

Tout compact Ω contenant $\sigma_\varepsilon(A)$ est K-spectral pour A avec $K = \dfrac{\text{length}(\partial\Omega)}{2\pi \cdot \varepsilon}$.

Preuve On utilise la formule de Cauchy 1.1.11

$$\|f(A)\| = \left\| \frac{1}{2\pi i} \int_{\partial\Omega} f(w)\,(wI-A)^{-1}\,dw \right\|$$

$$\leq \frac{1}{2\pi} \int_{\partial\Omega} \underbrace{|f(w)|}_{\leq \|f\|_\Omega} \cdot \underbrace{\|(wI-A)^{-1}\|}_{\leq \frac{1}{\varepsilon} \text{ car } w \notin \text{Int}(\sigma_\varepsilon(A))} |dw|$$

$$\leq \frac{\|f\|_\Omega}{2\pi\,\varepsilon} \cdot \underbrace{\int_{\partial\Omega} |dw|}_{= \text{length}(\partial\Omega)} \qquad \blacksquare$$

On va expérimenter aux TP avec le pseudospectre.

1.4.12 Exercice :

Pour $\varepsilon > 0$, montrer que

$$\bigcup_{\|E\| \leq \varepsilon} W(A+E) = \{ z \in \mathbb{C} : \text{dist}(z, W(A)) \leq \varepsilon \}.$$