Séance 1

IV Recherche de valeurs propres

Dans cette partie on considère $A \in \mathcal{M}_n(\mathbb{C})$ et on s'intéresse au calcul approché de (certaines de) ses valeurs propres. Nous allons étudier :

- une méthode partielle qui permet de calculer une valeur propre particulière de A et le vecteur propre correspondant ;
- une méthode globale qui permet de calculer toutes les valeurs propres de A mais pas les vecteurs propres.

A) Sensibilité d'un problème aux valeurs propres.

Soit ε un « petit » réel. Soit $A_\varepsilon = \begin{pmatrix} 0 & \cdots & 0 & \varepsilon \\ 1 & \ddots & & 0 \\ & \ddots & \ddots & \vdots \\ 0 & & 1 & 0 \end{pmatrix} \in \mathcal{M}_n(\mathbb{C})$. Ses valeurs propres ont toutes pour module $\varepsilon^{1/n}$.

Si n est grand elles sont donc très éloignées de 0, unique valeur propre de A_0.

Quantifions cet éloignement :

Th Soit $A \in \mathcal{M}_n(\mathbb{C})$ diagonalisable : $A = PDP^{-1}$ avec $D_{ij} = \lambda_i \delta_{ij}$: $D = \mathrm{diag}((\lambda_i)_{i=1}^n)$.
Soit $\|\cdot\|$ une norme matricielle telle que $\|\mathrm{diag}((\mu_i)_{i=1}^n)\| \leq \|\mu\|_\infty$ pour tout $\mu = (\mu_i)_{i=1}^n \in \mathbb{C}^n$. On a $\sigma(A+\delta A) \subset \bigcup_{i=1}^n D_i$

où $D_i = \{ a \in \mathbb{C} \,/\, |a - \lambda_i| \leq \mathrm{cond}_{\|\cdot\|}(P) \, \|\delta A\| \}$

pour toute matrice $\delta A \in \mathcal{M}_n(\mathbb{C})$.

Rem Toutes les normes subordonnées $|||\cdot|||_p$ vérifient $|||\mathrm{diag}(\mu_i)|||_p \leq \|\mu\|_\infty$.

Démo a est valeur propre de $A+\delta A$ ssi $A+\delta A - aI$ n'est pas inversible,
i.e. ssi $P^{-1}(A+\delta A - aI)P$ n'est pas inversible,
i.e. ssi $D + P^{-1}\delta A P - aI$ n'est pas inversible

Scanned with CamScanner

① Si $D - aI$ n'est pas inversible, $\exists i \mid a = \lambda_i$ donc $a \in D_i$.

② Si $D - aI$ est inversible, a est valeur propre de $A + \delta A$ ssi

$$I + (D-aI)^{-1} P^{-1} \delta A P \text{ est non inversible},$$

ce qui implique que $\|(D-aI)^{-1} P^{-1} \delta A P\| \geq 1$

d'après un théorème démontré précédemment dans ce cours (car $\|\cdot\|$ est supposée matricielle). Or

$$\|(D-aI)^{-1} P^{-1} \delta A P\| \leq \|(D-aI)^{-1}\| \operatorname{cond}_{\|\cdot\|}(P) \|\delta A\| \leq \max_i \frac{1}{|\lambda_i - a|} \operatorname{cond}_{\|\cdot\|}(P) \|\delta A\|.$$

Donc $\max_i \frac{1}{|\lambda_i - a|} \operatorname{cond}_{\|\cdot\|}(P) \|\delta A\| \geq 1$ et $\min_i |\lambda_i - a| \leq \operatorname{cond}_{\|\cdot\|}(P) \|\delta A\|$. CQFD.

Remarquer que c'est le conditionnement de la matrice de passage qui a un rôle.

B) Une méthode partielle : la méthode de la puissance

Cette méthode, dont le but est d'approcher le mode propre (λ_1, x_1) de A où λ_1 est la valeur propre de plus grand module, consiste en l'algorithme :

- soit $q^{(0)} \in \mathbb{C}^n \mid \|q^{(0)}\|_2 = 1$
- définir les suites $(\lambda^{(k)})_{k \in \mathbb{N}}$ de $\mathbb{C}$ et $(q^{(k)})_{k \in \mathbb{N}}$ de $\mathbb{C}^n$ par :

$$\begin{cases} z^{(k)} = A q^{(k-1)}, & k \geq 1 \\ q^{(k)} = \dfrac{z^{(k)}}{\|z^{(k)}\|_2} \\ \lambda^{(k)} = q^{(k)*} A q^{(k)} \end{cases}$$

Th Si A est diagonalisable, de modes propres $(\lambda_i, x_i)_{i=1}^n$ ordonnés de telle manière que $|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|$, si $q^{(0)} = \sum_{i=1}^n a_i x_i$ avec $a_1 \neq 0$, l'algorithme est bien défini et $\exists C \in \mathbb{R}$, $\exists (\eta^{(k)})_{k \in \mathbb{N}}$ suite de $\mathbb{C}$ avec $|\eta^{(k)}| = 1 \; \forall k$ tels que :

$$|\lambda^{(k)} - \lambda_1| \leq C \left|\frac{\lambda_2}{\lambda_1}\right|^k$$

$$\|\eta^{(k)} q^{(k)} - \tilde{x}_1\|_2 \leq C \left|\frac{\lambda_2}{\lambda_1}\right|^k \qquad \forall k$$

où $\tilde{x}_1 = \dfrac{x_1}{\|x_1\|_2}$.

Rem Sous les hypothèses du théorème la méthode converge...

– Dans la démonstration on suppose que $\|x_1\|_2 = 1$ ($x_1 = \tilde{x}_1$).

<u>Démo</u> Remarquons une fois pour toutes que $\|q^{(k)}\|_2 = 1$. Ceci va nous permettre d'écrire $q^{(k)}$, combinaison linéaire de vecteurs, comme une certaine expression vectorielle divisée par la norme de cette expression, systématiquement. On a $q^{(k)} = \dfrac{A^k q^{(0)}}{\|A^k q^{(0)}\|_2}$. Puisque $q^{(0)} = \sum a_i x_i$ (cette expression en fonction des x_i est possible pour tout $q^{(0)}$ puisque A est diagonalisable),

$$q^{(k)} = \frac{\sum_i a_i \lambda_i^k x_i}{\|\sum_i a_i \lambda_i^k x_i\|_2} = \frac{\sum_i a_i \left(\frac{\lambda_i}{|\lambda_1|}\right)^k x_i}{\left\|\sum_i a_i \left(\frac{\lambda_i}{|\lambda_1|}\right)^k x_i\right\|_2}$$

$$= \frac{a_1 \left(\frac{\lambda_1}{|\lambda_1|}\right)^k x_1 + \sum_{i\geq 2} a_i \left(\frac{\lambda_i}{|\lambda_1|}\right)^k x_i}{\left\|a_1 \left(\frac{\lambda_1}{|\lambda_1|}\right)^k x_1 + \sum_{i\geq 2} a_i \left(\frac{\lambda_i}{|\lambda_1|}\right)^k x_i\right\|_2}$$

Notons $\varepsilon = \dfrac{\lambda_1}{|\lambda_1|} \in \mathbb{C}$. On a

$$\overline{\varepsilon^k}\, q^{(k)} = \frac{a_1 x_1 + \sum_{i\geq 2} a_i \left(\frac{\lambda_i}{|\lambda_1|}\right)^k \overline{\varepsilon^k}\, x_i}{\left\|a_1 x_1 + \sum_{i\geq 2} a_i \left(\frac{\lambda_i}{|\lambda_1|}\right)^k \overline{\varepsilon^k}\, x_i\right\|_2}.$$

Notons $\alpha = \dfrac{a_1}{|a_1|} \in \mathbb{C}$ ($a_1 \neq 0$). On a

$$\overline{\alpha}\, \overline{\varepsilon^k}\, q^{(k)} = \frac{|a_1| x_1 + \sum_{i\geq 2} \overline{\varepsilon^k}\, \overline{\alpha}\, a_i \left(\frac{\lambda_i}{|\lambda_1|}\right)^k x_i}{\left\||a_1| x_1 + \sum_{i\geq 2} \overline{\varepsilon^k}\, \overline{\alpha}\, a_i \left(\frac{\lambda_i}{|\lambda_1|}\right)^k x_i\right\|_2}.$$

Notons $\eta^{(k)} = \overline{\alpha}\, \overline{\varepsilon^k}$. On a $|\eta^k| = 1$ et

$$\eta^{(k)} q^{(k)} - x_1 = \left(\frac{|a_1|}{\left\||a_1| x_1 + \sum_{i\geq 2} \eta^{(k)} a_i \left(\frac{\lambda_i}{|\lambda_1|}\right)^k x_i\right\|_2} - 1\right) x_1 + \frac{\sum_{i\geq 2} \eta^{(k)} a_i \left(\frac{\lambda_i}{|\lambda_1|}\right)^k x_i}{\left\||a_1| x_1 + \sum_{i\geq 2} \eta^{(k)} a_i \left(\frac{\lambda_i}{|\lambda_1|}\right)^k x_i\right\|_2}$$

Scanned with CamScanner

On va montrer que les deux termes du membre de droite sont petits...

Notons $\delta^{(k)} = \sum_{i\geq 2} \eta^{(k)} a_i \left(\frac{\lambda_i}{|\lambda_1|}\right)^k x_i$. Il existe $\tilde{C} \,/\, \|\delta^{(k)}\|_2 \leq \tilde{C}\left|\frac{\lambda_2}{\lambda_1}\right|^k$

(car $|\lambda_i| \geq |\lambda_2| \;\; \forall i \geq 2$). D'autre part,

$$\| |a_1| x_1 + \delta^{(k)} \|_2 \geq \| |a_1| x_1 \|_2 - \| \delta^k \|_2 \quad \text{(inégalité triangulaire)}$$

$$\geq \frac{1}{2} \| |a_1| x_1 \|_2 \quad \text{pour } k \text{ assez grand, d'après}$$

Donc $$\|\eta^{(k)} q^{(k)} - x_1\|_2 = \left\| \left(\frac{|a_1|}{\| |a_1| x_1 + \delta^{(k)} \|_2} - 1 \right) x_1 + \frac{\delta^{(k)}}{\| |a_1| x_1 + \delta^{(k)} \|_2} \right\|_2$$

$$\leq \left| \frac{|a_1|}{\| |a_1| x_1 + \delta^{(k)} \|_2} - 1 \right| + \frac{\|\delta^{(k)}\|_2}{\| |a_1| x_1 + \delta^{(k)} \|_2}$$

$$\leq \frac{\left| |a_1| - \| |a_1| x_1 + \delta^{(k)} \|_2 \right|}{\| |a_1| x_1 + \delta^{(k)} \|_2} + \frac{\|\delta^{(k)}\|_2}{\frac{1}{2} \| |a_1| x_1 \|_2} \quad \text{(noter que } \| |a_1| x_1 \|_2 = |a_1| \text{)}$$

$$\leq \frac{\| |a_1| x_1 - (|a_1| x_1 + \delta^{(k)}) \|_2}{\| |a_1| x_1 + \delta^{(k)} \|_2} + \frac{2\tilde{C}}{|a_1|} \left| \frac{\lambda_2}{\lambda_1} \right|^k$$

$$\leq \frac{\|\delta^{(k)}\|_2}{\| |a_1| x_1 + \delta^{(k)} \|_2} + \frac{2\tilde{C}}{|a_1|} \left| \frac{\lambda_2}{\lambda_1} \right|^k \leq \frac{4\tilde{C}}{|a_1|} \left| \frac{\lambda_2}{\lambda_1} \right|^2.$$ Il suffit

de poser $C = \frac{4\tilde{C}}{|a_1|}$ pour obtenir la deuxième inégalité énoncée.

Scanned with CamScanner

Par ailleurs, $\lambda^{(k)} = q^{(k)*} A\, q^{(k)} = \left(\eta^{(k)} q^{(k)}\right)^* A \left(\eta^{(k)} q^{(k)}\right)$ car $|\eta^{(k)}| = 1$.

Donc $\lambda^{(k)} = \left(x_1 + \eta^{(k)} q^{(k)} - x_1\right)^* A \left(x_1 + \eta^{(k)} q^{(k)} - x_1\right)$

$$= x_1^* A x_1 + x_1^* A \left(\eta^{(k)} q^{(k)} - x_1\right) + \left(\eta^{(k)} q^{(k)} - x_1\right)^* A x_1$$

$$= \lambda_1 + x_1^* A \left(\eta^{(k)} q^{(k)} - x_1\right) + \left(\eta^{(k)} q^{(k)} - x_1\right)^* A x_1$$

et $|\lambda^{(k)} - \lambda_1| \leq \left|x_1^* A \left(\eta^{(k)} q^{(k)} - x_1\right)\right| + \left|\left(\eta^{(k)} q^{(k)} - x_1\right)^* A x_1\right|$

$\leq D \left|\frac{\lambda_2}{\lambda_1}\right|^k$ d'après ce qu'on a montré juste avant, pour un certain D.

Finalement, on a donc les 2 inégalités de l'énoncé avec
$$C = \max\left(D, \frac{4\tilde{C}}{|\alpha_1|}\right).$$

Question : quand arrêter l'algorithme ?

On peut stopper lorsque $\|q^{(k)} - q^{(k-1)}\|_2$ est inférieur à une tolérance donnée (ou $|\lambda^{(k)} - \lambda^{(k-1)}|$ ...), mais un test plus fin consiste à évaluer $Aq^{(k)} - \lambda^{(k)} q^{(k)}$, que l'on note $r^{(k)}$ et que l'on appelle le résidu à l'étape k. D'après le théorème que nous venons de démontrer, $r^{(k)} \xrightarrow[k\to\infty]{} 0$ (à vérifier à titre d'exercice). L'analyse qui suit permet, au moyen du résidu, de calculer une bonne estimation de l'erreur faite à chaque étape (à suivre).

Scanned with CamScanner

Séance 2

Avec le résidu $r^{(k)} = A q^{(k)} - \lambda^{(k)} q^{(k)}$, définissons la matrice

$$E^{(k)} = - \frac{r^{(k)} q^{(k)*}}{\|r^{(k)}\|_2}.$$ C'est une matrice de rang 1 (exercice).

Elle vérifie $|||E^{(k)}|||_2 = 1$. En effet, pour tout $u \in \mathbb{C}^n$,

$$\|E^{(k)} u\|_2 = \frac{1}{\|r^{(k)}\|_2} \|(r^{(k)} q^{(k)*}) u\|_2 = \frac{1}{\|r^{(k)}\|_2} \|r^{(k)} (q^{(k)*} u)\|_2$$

$$= |q^* u| \leq \|q^{(k)}\|_2 \|u\|$$ (inégalité de Cauchy-Schwarz),

donc $|||E^{(k)}|||_2 \leq \|q^{(k)}\|_2 = 1$,

et, d'autre part, $\|E^{(k)} q^{(k)}\|_2 = \| q^{(k)*} q^{(k)} \| = 1$, donc $|||E^{(k)}|||_2 \geq 1$.

On a $\left(A + \|r^{(k)}\|_2 E^{(k)}\right) q^{(k)} = A q^{(k)} - r^{(k)} q^{(k)*} q^{(k)} = r^{(k)} + \lambda^{(k)} q^{(k)} - r^{(k)}$

$$= \lambda^{(k)} q^{(k)}.$$

Donc $(\lambda^{(k)}, q^{(k)})$ est un mode propre de la matrice $A + \|r^{(k)}\|_2 E^{(k)}$

(que l'on peut voir comme une perturbation de A) !

Existe-t-il un lien entre les valeurs propres de A et celles de ses perturbations ? Oui !

__Th__ Soit $A \in \mathcal{M}_n(\mathbb{C})$ diagonalisable. Soit λ une valeur propre simple de A, soit $x \in \mathbb{C}^n \setminus \{0\}$ tel que $Ax = \lambda x$, soit $y \in \mathbb{C}^n \setminus \{0\}$ tel que $A^* y = \bar{\lambda} y$. Soit $E \in \mathcal{M}_n(\mathbb{C})$ tel que $|||E|||_2 = 1$, soit $\varepsilon \in \mathbb{R}$: définissons $A(\varepsilon) = A + \varepsilon E$.

Il existe une fonction de classe C^1 : $V(0) \longrightarrow \mathbb{C} \times \mathbb{C}^n$ où $V(0)$ est un voisinage de 0

$$\varepsilon \longmapsto (\lambda(\varepsilon), x(\varepsilon))$$

telle que $\lambda(0) = \lambda$, $x(0) = x$, $A(\varepsilon) x(\varepsilon) = \lambda(\varepsilon) x(\varepsilon)$ $\forall \varepsilon \in V(0)$, et $\left|\frac{\partial \lambda}{\partial \varepsilon}(0)\right| \leq \frac{1}{|y^* x|}$

Scanned with CamScanner

Remarque Si λ est valeur propre simple, si x est vecteur propre à droite associé à λ, si y^* est vecteur propre à gauche associé à λ ($y^* A = \lambda y^*$, c'est-à-dire $A^* y = \bar{\lambda} y$), alors $y^* x \neq 0$, donc la majoration à la fin de l'énoncé a un sens (exercice : démontrer cette assertion)(A est diagonalisable).

Nous admettons ce théorème.

D'après la majoration sur la dérivée de la valeur propre $\lambda(\varepsilon)$ par rapport à ε, on a : $|\lambda(\varepsilon) - \lambda| \leqslant \dfrac{|\varepsilon|}{y^* x} + o(\varepsilon)$.

En utilisant ce théorème à l'étape k de la méthode de la puissance ($E^{(k)}$ joue le rôle de E, et $\|r^{(k)}\|_2$ celui de ε), on a : $|\lambda^{(k)} - \lambda_1| \leqslant \dfrac{\|r^{(k)}\|_2}{y_1^* x_1} + o(\|r^{(k)}\|_2)$ où λ_1 est la valeur propre cherchée, et x_1, y_1^* les vecteurs propres à droite et à gauche associés (pour A).

Attention : en réalité on ne connait pas x_1 ni y_1^* : on cherche justement à les approcher. Cependant cette majoration est un bon moyen d'estimer $|\lambda^{(k)} - \lambda_1|$ en remplaçant, dans le terme de droite, $\dfrac{\|r^{(k)}\|_2}{y_1^* x_1}$ par $\dfrac{\|r^{(k)}\|_2}{p^{(k)*} q^{(k)}}$ où $p^{(k)}$ est une approximation de y_1. Pour cela il faut un peu modifier l'algorithme de

Scanned with CamScanner

la puissance (afin qu'il fournisse cette approximation du vecteur propre à gauche). Voici cet algorithme :

- on se donne $q^{(0)} \in \mathbb{C}^n$ tel que $\|q^0\|_2 = 1$, et $p^0 \,/\, \|p^0\|_2 = 1$
- on définit les suites (λ^k), $(p^{(k)})$ et $(q^{(k)})$ par :

$$\begin{cases} z^{(k)} = A\, q^{k-1} \quad \text{et} \quad w^{(k)} = A^* \, p^{(k-1)} & \text{pour } k \geq 1, \\ q^{(k)} = \dfrac{z^{(k)}}{\|z^{(k)}\|_2} \quad \text{et} \quad p^{(k)} = \dfrac{w^{(k)}}{\|w^{(k)}\|_2} & \\ \lambda^{(k)} = q^{(k)*} A\, q^{(k)} & \\ e^{(k-1)} = \dfrac{\|z^{(k)} - \lambda^{(k-1)} q^{(k-1)}\|_2}{p^{(k-1)*}\, q^{(k-1)}} = \dfrac{\|r^{(k-1)}\|_2}{p^{(k-1)*}\, q^{(k-1)}} & \end{cases}.$$

Le réel $e^{(k-1)}$ fournit une majoration (approchée !) de $|\lambda^{(k-1)} - \lambda_1|$: on peut décider de stopper l'algorithme lorsque $e^{(k-1)} \leq \delta$ où δ est une « tolérance » donnée (par exemple, 10^{-6} ...).

La méthode de la puissance permet aussi (dans certains cas) de calculer une approximation du mode propre associé à la valeur propre de plus petit module de A (si cette valeur propre est non nulle) : soit A une matrice inversible et diagonalisable. Notons $(\lambda_i, x_i)_{i=1}^{n}$ ses modes propres, et supposons

$$0 < |\lambda_1| < |\lambda_2| \leq |\lambda_3| \leq \cdots \leq |\lambda_n|.$$

A est inversible

Alors la méthode de la puissance appliquée à A^{-1} permet le calcul approché de (λ_1, x_1). Bien entendu, il ne s'agit en aucun cas de calculer A^{-1} pour mettre en œuvre l'algorithme, mais de résoudre des systèmes associés à A :

l'équation $z^{(k)} = A^{-1} q^{(k-1)}$ est en fait

$A z^{(k)} = q^{(k-1)}$, qui peut se résoudre au moyen de méthodes étudiées dans la partie III du cours par exemple.

Cette méthode dédiée au calcul approché de (λ_1, x_1) s'appelle méthode de la puissance inverse.

Exercice : écrire complètement l'algorithme de la puissance inverse.

c) Une méthode globale : la méthode QR.

Soit $A \in \mathcal{M}_n(\mathbb{C})$ de modes propres $(\lambda_i, x_i)_{i=1}^{n}$.

On suppose qu'$\exists$ $p \leq n$ tel que $|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_p| > |\lambda_{p+1}| \geq \dots \geq |\lambda_n| > 0$

A admet une décomposition de Schur : $A = Q T Q^*$ avec Q unitaire et T triangulaire supérieure.

Notons $Q = \left(Q_1 \middle| Q_2 \right)$ où $Q_1 \in \mathcal{M}_{n,p}(\mathbb{C})$ et $Q_2 \in \mathcal{M}_{n,n-p}(\mathbb{C})$: les colonnes de Q_1 sont les p premières colonnes de Q, celles de Q_2 sont les $(n-p)$ dernières.

Scanned with CamScanner

Nous allons chercher à approcher le sous-espace vectoriel (de dimension p) engendré par les colonnes de Q_1, sous-espace vectoriel que l'on note $D_p(A)$ (cette approximation de $D_p(A)$ est ici rendue possible par l'hypothèse $|\lambda_p| > |\lambda_{p+1}|$).

Attention $D_p(A)$ n'est pas le sous-espace vectoriel engendré par les $(x_i)_{i=1}^{p}$, sauf si T est diagonale (en particulier, sauf si A est hermitienne et donc diagonalisable en base orthonormée).

L'algorithme que nous allons utiliser, appelé algorithme QR, est une généralisation, une extension de l'algorithme de la puissance. Il consiste à

- se donner $Q^{(0)} \in \mathcal{M}_{n,p}(\mathbb{C})$ dont les colonnes sont orthonormées
- définir les suites $(R^{(k)})_{k\geq 1}$ et $(Q^{(k)})_{k\geq 1}$ par

$$\left\{\begin{array}{l} Z^{(k)} = A\,Q^{(k-1)} \\ Z^{(k)} = Q^{(k)} R^{(k)} \quad \text{où } Q^{(k)} \in \mathcal{M}_{n,p}(\mathbb{C}) \text{ a} \\ \qquad \text{ses colonnes orthonormées} \\ \text{et } R^{(k)} \in \mathcal{M}_p(\mathbb{C}) \text{ est triangulaire supérieure} \end{array}\right.$$

Dans l'équation $Z^{(k)} = Q^{(k)} R^{(k)}$, ce sont donc $Q^{(k)}$ et $R^{(k)}$ les inconnues à calculer : $Q^{(k)} R^{(k)}$ est une factorisation QR de $Z^{(k)}$. On peut utiliser l'algorithme de Householder, par exemple, pour les calculer

Scanned with CamScanner

Remarque si $p = 1$, l'hypothèse $|\lambda_p| > |\lambda_{p+1}|$ est celle que l'on a faite pour l'algorithme de la puissance, et l'on remarque que l'algorithme QR est exactement l'algorithme de la puissance, et l'équation $Z^{(k)} = Q^{(k)} R^{(k)}$ (où $R^{(k)} \in \mathcal{M}_1(\mathbb{C}) \equiv \mathbb{C}$) est une normalisation de $Z^{(k)}$ (en $Q^{(k)}$).

Th Avec les notations introduites en début de chapitre, et sous l'hypothèse $|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_p| > |\lambda_{p+1}| \geq \ldots \geq |\lambda_n| > 0$, si $\operatorname{dist}\left(D_p(A^*), \operatorname{Im} Q^{(0)}\right) < 1$ on a

$$\operatorname{dist}\left(D_p(A), \operatorname{Im} Q^{(k)}\right) = \mathcal{O}\left(\left|\frac{\lambda_{p+1}}{\lambda_p}\right|^k\right).$$

Cet énoncé utilise la

Définition Soient E et F deux sous-espaces vectoriels de $\mathbb{C}^n$. On note $\operatorname{dist}(E, F) = |||P_E - P_F|||_2$ où P_E et P_F sont les projections orthogonales sur E et F.

Interprétation du théorème : si k est grand, les projection orthogonales sur $D_p(A)$ et sur $\operatorname{Im} Q^{(k)}$ (espace vectoriel engendré par les colonnes de $Q^{(k)}$) sont proches, $D_p(A)$ et $\operatorname{Im} Q^{(k)}$ sont proches...

Nous admettons ce théorème.

Scanned with CamScanner

Remarque sur l'hypothèse $q^{(0)} = \sum_{i=1}^{n} a_i x_i$ avec $a_1 \neq 0$ pour la convergence de la méthode de la puissance.

$a_1 \neq 0$ s'écrit aussi : $y_1^* q^{(0)} \neq 0$, où y_1^* est vecteur propre à gauche de A, associé à λ_1.

En effet : $y_1^* A = \lambda_1 y_1^*$, soit $A^* y_1 = \bar{\lambda}_1 y_1$.

Donc $y_1^* A x_i = \lambda_i y_1^* x_i \quad \forall i = 1, \dots, n$.

$$\begin{matrix} y_1^* A x_i \\ \parallel \\ \lambda_1 y_1^* x_i \end{matrix}$$

Comme $\lambda_1 \neq \lambda_i$ si $i \neq 1$, pour $i \neq 1$ on a donc $y_1^* x_i = 0$.

Donc $y_1^* x_1 \neq 0$, sinon y_1^* serait nul (car $(x_i)_{i=1}^{n}$ est une base de $\mathbb{C}^n$, hypothèse que nous avons faite pour la méthode de la puissance).

cf. remarque en haut de la p. 7

Ainsi $y_1^* q^{(0)} = \sum_i a_i y_1^* x_i = a_1 y_1^* x_1$:

$y_1^* q^{(0)} \neq 0 \iff a_1 \neq 0$.

On dit que $y_1^* q^{(0)} \neq 0$ revient à dire que

$\mathrm{dist}\left(\mathrm{vect}(y_1), \mathrm{vect}(q^{(0)})\right) < 1$ (dans le cas A diagonalisable).

On comprend que les hypothèses sur la donnée initiale dans l'algorithme de la puissance et dans l'algorithme QR sont proches.

Scanned with CamScanner

<u>Remarque</u> Dans le cas $p=1$ toujours, ce théorème (admis) donne la convergence de la méthode de la puissance sans l'hypothèse de « diagonalisabilité » de A : il est plus fort que celui que nous avons démontré. La seule hypothèse réellement nécessaire sur A pour la convergence de la méthode de la puissance est que $|\lambda_1| > |\lambda_2|$ ...

(À suivre)

Scanned with CamScanner

<u>Application</u> dans le cas particulier où $|\lambda_1| > |\lambda_2| > \dots > |\lambda_n| > 0$ (pour toute la suite).

Remarquons que dans ce cas A a n valeurs propres distinctes, A est donc diagonalisable. Plus précisément : soient $(\lambda_i, x_i)_{i=1}^{n}$ les modes propres de A. Soient $(y_i^*)_{i=1}^{n}$ les vecteurs propres à gauche (associés aux λ_i) de A : $y_i^* A = \lambda_i y_i^* \quad \forall i$, soit

$A^* y_i = \bar{\lambda}_i y_i$: y_i est vecteur propre (à droite) de A^*, associé à la valeur propre $\bar{\lambda}_i$ de A^* (pour tout i).

Pour tout $(i,j) \in \{1,\dots,n\}^2$ on a $y_j^* A x_i = \lambda_j y_j^* x_i = \lambda_i y_j^* x_i$.

Donc, si $i \neq j$, $y_j^* x_i = 0$. Pour j donné, y_j est orthogonal à tous les x_i pour $i \neq j$.

Dans le cas où A est hermitienne, les x_i sont orthogonaux, les y_i sont orthogonaux et $y_i = x_i \quad \forall i$ (à une normalisation près)

On applique l'algorithme avec $p = n$ (c'est-à-dire que l'on choisit $Q^{(0)} \in \mathcal{M}_n(\mathbb{C})$).
Comme on peut appliquer le résultat (du théorème) pour tout $p \in \{1,\dots,n\}$,

si $\operatorname{dist}\left(D_i(A), \operatorname{vect}\left((q_j^{(0)})_{j=1}^{i}\right)\right) < 1 \quad \forall i$, on a

$$\operatorname{dist}\left(D_i(A), \operatorname{vect}\left((q_j^{(k)})_{j=1}^{i}\right)\right) \xrightarrow[k\to\infty]{} 0 \quad \forall i.$$

Scanned with CamScanner

Ceci implique que $T^{(k)} = Q^{(k)*} A\, Q^{(k)}$ converge vers une factorisation de Schur de A : $T^{(k)}$ tend vers une matrice triangulaire, avec des coefficients diagonaux qui tendent vers les valeurs propres de A.

<u>En effet</u> : • $T = Q^* A Q$ avec T triangulaire supérieure s'écrit :

$AQ = QT$, soit $A q_i = Q t_i \quad \forall i$, où les q_i et t_i sont les colonnes de Q et T. Puisque T est triangulaire supérieure, avec les valeurs propres λ_i de A sur la diagonale, ceci s'écrit encore

$$\begin{cases} A q_1 = \lambda_1 q_1 \\ A q_2 = T_{12}\, q_1 + \lambda_2 q_2 \\ A q_3 = T_{13}\, q_1 + T_{23}\, q_2 + \lambda_3 q_3 \\ \quad \vdots \\ A q_i = T_{1i}\, q_1 + T_{2i}\, q_2 + \dots + \lambda_i q_i \\ \quad \vdots \end{cases}$$

• Pour $i=1$: $\text{dist}\left(D_1(A), \text{vect}(q_1^{(k)})\right) \longrightarrow 0$ implique qu'il existe $(\eta^{(k)})_{k\in\mathbb{N}}$, suite de complexes de module 1, telle que $\eta^{(k)} q_1^{(k)} \longrightarrow q_1$, donc

$$\eta^{(k)} A\, q_1^{(k)} = \eta^{(k)} Q^{(k)} t_1^{(k)} = \eta^{(k)} \Big(T_{11}^{(k)} q_1^{(k)} + \sum_{j>1} T_{j1}^{(k)} q_j^{(k)} \Big)$$

($t_1^{(k)}$ désigne la première colonne de T^k)

$\downarrow$

$\lambda_1 q_1$

implique bien que $T_{11}^{(k)} \longrightarrow \lambda_1$

et que $T_{j1}^{(k)} \longrightarrow 0 \quad \forall j > 1$.

• Pour $i = 2$: compte tenu du point ci-dessus pour $i=1$, le fait que $\text{dist}\left(D_2(A), \text{vect}(q_1^{(k)}, q_2^{(k)})\right) \xrightarrow[k\to\infty]{} 0$ implique que

Scanné avec CamScanner

$q_2^{(k)}$ est proche (pour k grand) d'une combinaison linéaire de q_1 et q_2 : $\exists\ (\eta_1^{(k)})$ et $(\eta_2^{(k)})$ telles que

$$\begin{cases} q_2^{(k)} - \left(\eta_1^{(k)} q_1 + \eta_2^{(k)} q_2\right) \xrightarrow[k\to\infty]{} 0 \\ \text{et } \|\eta_1^{(k)} q_1 + \eta_2^{(k)} q_2\|_2 = 1 \end{cases}$$

(Le point de vue est ici un peu différent de celui du point précédent : on pourrait réécrire le point $i=1$ sous la forme

$$q_1^{(k)} - \eta^{(k)} q_1 \longrightarrow 0 \ldots)$$

Comme $Q^{(k)}$ est une matrice orthogonale (unitaire), les vecteurs $(q_\ell^{(k)})_{\ell=1}^{m}$ sont orthogonaux. La convergence $q_2^{(k)} - (\eta_1^{(k)} q_1 + \eta_2^{(k)} q_2) \longrightarrow 0$ implique que $q_\ell^{(k)*} A\, q_2^{(k)} \xrightarrow[k\to\infty]{} 0 \quad \forall \ell \geq 3$: c'est-à-dire que $T_{\ell 2}^{(k)} \longrightarrow 0 \quad \forall \ell \geq 3$.

- pour i quelconque : une récurrence basée sur le principe vu ci-dessus permet de conclure...

De conclure quoi ? Que $T^{(k)} = Q^{(k)*} A\, Q^{(k)}$ a ses coefficients sous diagonaux qui convergent vers 0.

Scanned with CamScanner

Remarques très importantes.

a. Attention : $T^{(k)}$ est « presque » triangulaire lorsque k est grand, mais pas presque diagonale, alors que A est diagonalisable, a priori : les colonnes de $Q^{(k)}$ ne convergent pas (a priori) vers les vecteurs propres de A.

Il existe cependant un cas particulier intéressant : celui où A est hermitienne. Dans ce cas en effet, si

$A = QTQ^*$ est une factorisation de Schur de A

(T est triangulaire supérieure et Q est unitaire),

$A^* = QT^*Q^* = A$ implique que $T = T^*$, donc T est non seulement triangulaire mais encore diagonale (toute factorisation de Schur d'une matrice hermitienne est une diagonalisation).

Si A est hermitienne donc, $T^{(k)}$ converge vers une matrice diagonale.

b. Dans le cas général, si A n'est pas hermitienne, en particulier, la méthode QR ne permet pas d'approcher les vecteurs propres de A (sauf x_1). Pour ce faire, une méthode souvent utilisée est

① de calculer, à l'aide de la méthode QR, des approximations $\tilde{\lambda}_i$ pas trop précises des λ_i

② d'approcher précisément les λ_i et les x_i à l'aide de la méthode de la puissance inverse appliquée à $A - \tilde{\lambda}_i I$: les $\tilde{\lambda}_i$ doivent

Scanned with CamScanner

donc n'être pas trop proches des λ_i, car $A-\lambda_i I$ est singulière !

c. Il existe une autre façon d'écrire l'algorithme de la méthode QR, nouvelle écriture facile à retenir.

Chaque itération a jusqu'ici été écrite

$$\begin{cases} Z^{(k)} = A Q^{(k-1)} \\ Z^{(k)} = Q^{(k)} R^{(k)} \end{cases} \quad \text{(factorisation QR de } Z^{(k)}\text{)}.$$

En posant $T^{(k)} = Q^{(k)*} A Q^{(k)}$ pour tout k, on a

$$T^{(k-1)} = Q^{(k-1)*} A Q^{(k-1)} = Q^{(k-1)*} Z^{(k)} = Q^{(k-1)*} Q^{(k)} R^{(k)}.$$

remarque que $U^{(k)}$ est unitaire $\longrightarrow$ Notons $U^{(k)} = Q^{(k-1)*} Q^{(k)}$: on a $T^{(k-1)} = U^{(k)} R^{(k)}$.

Or $Q^{(k)} = Q^{(k-1)} U^{(k)}$, donc on a aussi

$$T^{(k)} = Q^{(k)*} A Q^{(k)} = Q^{(k)*} A Q^{(k-1)} U^{(k)} = Q^{(k)*} Z^{(k)} U^{(k)}$$
$$= Q^{(k)*} Q^{(k)} R^{(k)} U^{(k)} = R^{(k)} U^{(k)}.$$

Donc $T^{(k-1)} = U^{(k)} R^{(k)}$ et $T^{(k)} = R^{(k)} U^{(k)}$.

L'algorithme se réécrit

- Choisir $Q^{(0)}$ unitaire (par exemple, $Q^{(0)} = I$).
- Poser $T^{(0)} = Q^{(0)*} A Q^{(0)}$
- Pour $k \geq 1$, poser $\begin{cases} T^{(k-1)} = U^{(k)} R^{(k)} \quad \text{(factorisation QR de } T^{(k-1)}\text{)} \\ T^{(k)} = R^{(k)} U^{k} \end{cases}$.

d. L'étude de la méthode QR est très délicate, c'est pourquoi nous ne l'avons pas faite (et avons admis un gros résultat).

Le fait que la diagonale de $T^{(k)}$ converge vers un vecteur contenant toutes les valeurs propres de A est vrai <u>sans</u> l'hypothèse $\mathrm{dist}\left(D_i(A), \mathrm{vect}\left((q_j^{(0)})_{j=1}^{i}\right)\right) < 1$ pour tout i, mais dans ce cas on n'a aucune information sur l'<u>ordre</u> dans lequel ces valeurs propres apparaissent. Pour s'en convaincre, il suffit de considérer une matrice A triangulaire (avec des valeurs propres positives) et de constater qu'alors l'algorithme QR donne(*), pour tout k, $T^{(k)} = A$ (si la factorisation QR choisie est celle où R a des coefficients positifs sur la diagonale).
Donc pour cette matrice A l'ordre des valeurs propres sur la diagonale n'est pas modifié au cours des itérations...

* Si l'on choisit $Q^{(0)} = I$.

Scanned with CamScanner