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 $u \in \mathbb{C}^n$,
$$\|E^{(k)} u\|_2 = \frac{1}{\|r^{(k)}\|_2} \|(r^{(k)} q^{(k)*}) u\|_2 = \frac{1}{\|r^{(k)}\|_2} \| r^{(k)} (q^{(k)*} u)\|_2$$
$$= |q^* u| \leq \|q^{(k)}\|_2 \|u\| \text{ (inégalité de Cauchy-Schwarz)},$$
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 $\left|\frac{\partial \lambda}{\partial \varepsilon}(0)\right| \leq \frac{1}{|y^* x|}$.

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 : $|\lambda(\varepsilon) - \lambda| \leq \frac{|\varepsilon|}{y^* x} + o(\varepsilon)$.

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 : $|\lambda^{(k)} - \lambda_1| \leq \frac{\|r^{(k)}\|_2}{y_1^* x_1} + o(\|r^{(k)}\|_2)$ 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, $\frac{\|r^{(k)}\|_2}{y_1^* x_1}$ par $\frac{\|r^{(k)}\|_2}{p^{(k)*} q^{(k)}}$ où $p^{(k)}$ est une approximation de y_1. Pour cela il faut un peu modifier l'algorithme de 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 :

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

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 <u>plus petit</u> module de A (si cette valeur propre est non nulle) : soit A une matrice <u>inversible</u> et diagonalisable. Notons $(\lambda_i, x_i)_{i=1}^n$ ses modes propres, et supposons

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

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'$\exists$ $p < n$ tel que $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_p| > |\lambda_{p+1}| \geq \cdots \geq |\lambda_n|$.

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

Notons $Q = (Q_1 | Q_2)$ 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.

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 $|\lambda_p| > |\lambda_{p+1}|$).

<u>Attention</u> $D_p(A)$ <u>n'est pas</u> 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

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

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.

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 $|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_p| > |\lambda_{p+1}| \geq \dots \geq |\lambda_n|$, 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 projections 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.

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

$$\underset{\lambda_1 y_1^* x_i}{\overset{\parallel}{y_1^* A x_i}}$$

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 $\operatorname{dist}(\operatorname{vect}(y_1), \operatorname{vect}(q^{(0)})) < 1$ (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.

<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)