En injectant ceci dans (4.4), on arrive à l'estimation recherchée. □

4.4.3 Exemple 3

On pose $\boldsymbol{p}_k := \mathrm{A}\boldsymbol{x}^{(k)} - \boldsymbol{b}$, et on prend $\boldsymbol{v}_k = \mathrm{A}^\top \boldsymbol{p}_k$, ainsi que $\boldsymbol{w}_k = \mathrm{A}\boldsymbol{v}_k = \mathrm{A}\mathrm{A}^\top \boldsymbol{p}_k$. La méthode de projection 1D correspondante s'écrit alors

$$\begin{cases} \boldsymbol{p}_k = \mathrm{A}\boldsymbol{x}^{(k)} - \boldsymbol{b} \\ \boldsymbol{v}_k = \mathrm{A}^\top \boldsymbol{p}_k \\ \alpha_k = |\boldsymbol{v}_k|_2^2 / |\mathrm{A}\boldsymbol{v}_k|_2^2 \\ \boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} - \alpha_k \boldsymbol{v}_k. \end{cases}$$

Ici on ne suppose rien sur la matrice A sinon qu'elle est inversible.

4.5 Méthodes de Krylov

Rappelons que si $\mathrm{A} \in \mathbb{R}^{n\times n}$ alors d'après le théorème de Cayley-Hamilton, il existe $c_j \in \mathbb{R}, j = 1 \ldots n$ tel que $\mathrm{A}^n + c_1\mathrm{A}^{n-1} + \cdots + c_{n-1}\mathrm{A} + c_n\mathrm{Id} = 0$. Si A est inversible, on a $c_n \neq 0$ de sorte que $-(\mathrm{A}^n + \cdots + c_{n-1}\mathrm{A})/c_n = \mathrm{Id}$. Dans le membre de gauche de cette dernière égalité on peut factoriser A, et on obtient

$$\begin{aligned} &\mathrm{A}\cdot \mathrm{Q}(\mathrm{A}) = \mathrm{Id} \\ &\text{avec}\quad \mathrm{Q}(X) := -\frac{1}{c_n}X^{n-1} - \frac{c_1}{c_n}X^{n-2} - \cdots - \frac{c_{n-2}}{c_n}X - \frac{c_{n-1}}{c_n} \end{aligned}$$

c'est-à-dire $\mathrm{A}^{-1} = \mathrm{Q}(\mathrm{A})$ où $\mathrm{Q}(X) \in \mathbb{P}_{n-1}[X]$ est un polynôme de degré au plus $n-1$. A présent si l'on s'intéresse à résoudre $\mathrm{A}\boldsymbol{x}_\star = \boldsymbol{b}$, en considérant un $\boldsymbol{x}^{(0)} \in \mathbb{R}^n$ quelconque, et en posant $\boldsymbol{r}_0 := \mathrm{A}\boldsymbol{x}^{(0)} - \boldsymbol{b}$, on a

$$\begin{cases} \mathrm{A}(\boldsymbol{x}_\star - \boldsymbol{x}^{(0)}) = -(\mathrm{A}\boldsymbol{x}^{(0)} - \boldsymbol{b}) = -\boldsymbol{r}_0 \\ \boldsymbol{x}_\star = \boldsymbol{x}^{(0)} - \mathrm{A}^{-1}\boldsymbol{r}_0 = \boldsymbol{x}^{(0)} - \mathrm{Q}(\mathrm{A})\boldsymbol{r}_0 \end{cases}$$

Il est alors naturel de rechercher une approximation de $\boldsymbol{x}_\star$ dans l'espace $\boldsymbol{x}_0 + \mathfrak{K}_k$, pour un certain $k \geq 1$, où $\mathfrak{K}_k \subset \mathbb{R}^n$, appelé espace de Krylov d'ordre k, est défini par

$$\mathfrak{K}_k = \mathrm{vect}\{\boldsymbol{r}_0, \mathrm{A}\boldsymbol{r}_0, \ldots, \mathrm{A}^{k-1}\boldsymbol{r}_0\}$$

Lemme 4.5.1.
Il existe un entier $k_0 \leq n$ critique tel que $\mathfrak{K}_0 \subset \mathfrak{K}_1 \subset \cdots \subset \mathfrak{K}_{k_0} = \mathfrak{K}_{k_0+1} = \cdots = \mathfrak{K}_n$ avec $\mathfrak{K}_j \neq \mathfrak{K}_{j+1}$ pour $j < k_0$.

Démo :
Tout d'abord, on sait d'après le théorème de Cayley-Hamilton, que $\mathrm{A}^n\boldsymbol{r}_0 \in \mathfrak{K}_n$ de sorte que $\mathfrak{K}_n = \mathfrak{K}_{n+1}$. Nous allons montrer que pour k satisfaisant $\mathfrak{K}_k = \mathfrak{K}_{k+1}$, on a $\mathfrak{K}_k = \mathfrak{K}_{k+p} \forall p \geq 0$. On aura alors montré qu'un k_0 tel que mentionné dans l'énoncé existe et que forcément $k_0 \leq n$. Il restera alors à prendre le plus petit de tels k_0.

Soit donc $k \leq n$ tel que $\mathfrak{K}_k = \mathfrak{K}_{k+1}$. Procédons par récurrence sur p pour démontrer la propriété souhaitée. Pour $p = 1$ on sait que $\mathfrak{K}_k = \mathfrak{K}_{k+p}$. Supposons maintenant que pour $p \geq 0$ on ait $\mathfrak{K}_k = \mathfrak{K}_{k+1} = \cdots = \mathfrak{K}_{k+p}$ et montrons que $\mathfrak{K}_k = \mathfrak{K}_{k+p+1}$. Soit $\boldsymbol{x} \in \mathfrak{K}_{k+p+1}$. Ce vecteur se décompose sous la forme $\boldsymbol{x} = \boldsymbol{y} + \alpha \mathrm{A}^{k+p}\boldsymbol{r}_0$ pour $\boldsymbol{y} \in \mathfrak{K}_{k+p} = \mathfrak{K}_k, \alpha \in \mathbb{R}$. Il suffit donc de démontrer que $\mathrm{A}^{k+p}\boldsymbol{r}_0 \in \mathfrak{K}_k$ pour conclure que $\boldsymbol{x} \in \mathfrak{K}_k$, et donc que $\mathfrak{K}_{k+p+1} \subset \mathfrak{K}_k$. Comme $\mathrm{A}^{k+p-1}\boldsymbol{r}_0 \in \mathfrak{K}_{k+p} = \mathfrak{K}_k$, il existe des coefficients $c_j \in \mathbb{R}, j = 0 \ldots k-1$ tels que $\mathrm{A}^{k+p-1}\boldsymbol{r}_0 = c_0\boldsymbol{r}_0 + c_1\mathrm{A}\boldsymbol{r}_0 + \cdots + c_k\mathrm{A}^k\boldsymbol{r}_0$. On a alors

$$
\begin{aligned}
\mathrm{A}^{k+p}\boldsymbol{r}_0 &= \mathrm{A}(\mathrm{A}^{k+p-1}\boldsymbol{r}_0) \\
&= \mathrm{A}(\sum_{j=0}^{k-1} c_j \mathrm{A}^j \boldsymbol{r}_0) = \sum_{j=0}^{k-1} c_j \mathrm{A}^{j+1}\boldsymbol{r}_0 = \sum_{j=1}^{k} c_{j-1}\mathrm{A}^j\boldsymbol{r}_0 \in \mathfrak{K}_{k+1} = \mathfrak{K}_k
\end{aligned}
$$

Ce qui conclue la preuve. □

4.5.1 La méthode du gradient conjugué

La méthode du gradient conjugué s'applique dans le cas où la matrice A est symétrique définie positive, ce que nous supposerons jusqu'à la fin de ce paragraphe. Elle consiste à calculer la suite $(\boldsymbol{x}_k)_{k\geq 0}$ définie par :

$$
\begin{cases}
\boldsymbol{x}^{(k)} \in \boldsymbol{x}^{(0)} + \mathfrak{K}_k \\
\boldsymbol{y}^\top(\mathrm{A}\boldsymbol{x}^{(k)} - \boldsymbol{b}) = 0 \quad \forall \boldsymbol{y} \in \mathfrak{K}_k.
\end{cases}
\tag{4.5}
$$

D'après ce qui précède, ces équations définissent $\boldsymbol{x}^{(k)}$ de manière unique pour chaque k. Montrons que pour k suffisament grand, on obtient la solution du système linéaire de départ.

Lemme 4.5.2.
Soit k_0 l'entier critique du lemme 4.5.1 pour lequel $\mathfrak{K}_{k_0} = \mathfrak{K}_{k_0+1}$. Alors on a $\mathrm{A}\boldsymbol{x}^{(k_0)} = \boldsymbol{b}$ de sorte que $\boldsymbol{x}^{(k_0)} = \boldsymbol{x}_\star$.

Démo :
On a

$$
\mathrm{A}\boldsymbol{x}^{(k_0)} - \boldsymbol{b} = \mathrm{A}\,\underbrace{(\boldsymbol{x}^{(k_0)} - \boldsymbol{x}^{(0)})}_{\in \mathfrak{K}_{k_0}} + \underbrace{\mathrm{A}\boldsymbol{x}^{(0)} - \boldsymbol{b}}_{\in \mathfrak{K}_{k_0}}
$$

Comme par ailleurs on a $\mathrm{A}(\mathfrak{K}_{k_0}) \subset \mathfrak{K}_{k_0+1} = \mathfrak{K}_{k_0}$, on en déduit que $\mathrm{A}\boldsymbol{x}^{(k_0)} - \boldsymbol{b} \in \mathfrak{K}_{k_0}$. On impose aussi par ailleurs $\boldsymbol{y}^\top(\mathrm{A}\boldsymbol{x}^{(k_0)} - \boldsymbol{b}) = 0$ pour tout $\boldsymbol{y} \in \mathfrak{K}_{k_0}$. En prenant $\boldsymbol{y} = \mathrm{A}\boldsymbol{x}^{(k_0)} - \boldsymbol{b}$, on en tire $|\mathrm{A}\boldsymbol{x}^{(k_0)} - \boldsymbol{b}|_2^2 = 0 \Rightarrow \mathrm{A}\boldsymbol{x}^{(k_0)} - \boldsymbol{b} = 0$. □

Théorème 4.5.1.
On considère une matrice symétrique définie positive $\mathrm{A} \in \mathbb{R}^{n\times n}$. On considère $\boldsymbol{x}^{(0)}, \boldsymbol{b} \in \mathbb{R}^n$ fixés, et on pose $\boldsymbol{r}_0 := \boldsymbol{p}_0 := \mathrm{A}\boldsymbol{x}^{(0)} - \boldsymbol{b}$. On définit alors les trois suites $(\boldsymbol{x}^{(k)})_{k\geq 0}, (\boldsymbol{r}_k)_{k\geq 0}, (\boldsymbol{p}_k)_{k\geq 0}$ par :

$$
\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} - \alpha_k \boldsymbol{p}_k \tag{4.6a}
$$

$$
\boldsymbol{r}_{k+1} = \boldsymbol{r}_k - \alpha_k \mathrm{A}\boldsymbol{p}_k \tag{4.6b}
$$

$$
\boldsymbol{p}_{k+1} = \boldsymbol{r}_{k+1} + \beta_k \boldsymbol{p}_k \tag{4.6c}
$$

$$
\textit{avec} \quad \alpha_k = \frac{|\boldsymbol{r}_k|_2^2}{|\boldsymbol{p}_k|_\mathrm{A}^2} \quad \textit{et} \quad \beta_k = \frac{|\boldsymbol{r}_{k+1}|_2^2}{|\boldsymbol{r}_k|_2^2}
$$

Alors la suite $\boldsymbol{x}^{(k)}$ construite selon la récurrence (4.6) coincide avec la suite des itérés de la méthode du gradient conjugué (4.5).

Démo :

D'après les relations de récurrence (4.6a) et (4.6b) ci-dessus, on a $\mathrm{A}\boldsymbol{x}^{(k+1)} - \boldsymbol{b} - \boldsymbol{r}_{k+1} = \mathrm{A}\boldsymbol{x}^{(k)} - \boldsymbol{b} - \boldsymbol{r}_k = \mathrm{A}\boldsymbol{x}^{(k-1)} - \boldsymbol{b} - \boldsymbol{r}_{k-1} = \cdots = \mathrm{A}\boldsymbol{x}^{(0)} - \boldsymbol{b} - \boldsymbol{r}_0 = 0$. On en déduit donc simplement

$$\boldsymbol{r}^{(k)} = \mathrm{A}\boldsymbol{x}^{(k)} - \boldsymbol{b} \quad \forall k \geq 0.$$

D'après (4.6b) et (4.6c), on a aussi $\boldsymbol{r}_k, \boldsymbol{p}_k \in \mathfrak{K}_{k+1} \Rightarrow \boldsymbol{r}_k, \mathrm{A}\boldsymbol{p}_k \in \mathfrak{K}_{k+2} \Rightarrow \boldsymbol{r}_{k+1} \in \mathfrak{K}_{k+2} \Rightarrow \boldsymbol{p}_{k+1} \in \mathfrak{K}_{k+2}$. Par récurrence on en déduit alors

$$\boldsymbol{r}_k, \boldsymbol{p}_k \in \mathfrak{K}_{k+1} \quad \forall k \geq 0.$$

On en tire également $\boldsymbol{x}^{(k)} \in \boldsymbol{x}^{(0)} + \mathfrak{K}_k \Rightarrow \boldsymbol{x}^{(k+1)} \in \boldsymbol{x}^{(0)} + \mathfrak{K}_{k+1}$. En raisonnant à nouveau par récurrence, on en déduit de même

$$\boldsymbol{x}^k \in \boldsymbol{x}^{(0)} + \mathfrak{K}_k \quad \forall k \geq 0.$$

Nous venons d'établir que la suite des $\boldsymbol{x}^{(k)}$ construits selon (4.6) vérifient la première équation de (4.5). Pour conclure, d'après il reste donc à démontrer qu'on a par ailleurs $\boldsymbol{y}^\top(\mathrm{A}\boldsymbol{x}^{(k)} - \boldsymbol{b}) = 0\ \forall \boldsymbol{y} \in \mathfrak{K}_k$, c'est-à-dire $\boldsymbol{y}^\top \boldsymbol{r}_k = 0\ \forall \boldsymbol{y} \in \mathfrak{K}_k$. Pour ceci nous allons démontrer par récurrence sur k qu'on a, lorsque $k \geq 1$,

$$\begin{cases} \boldsymbol{r}_j^\top \boldsymbol{r}_k = 0 \\ \boldsymbol{p}_j^\top \mathrm{A}\boldsymbol{p}_k = 0 \quad \forall j = 0 \ldots k-1. \end{cases} \tag{4.7}$$

Il n'y a rien à démontrer pour $k = 0$. Supposons donc que ce soit vrai pour k, et démontrons que c'est encore vrai pour $k+1$. En utilisant (4.6b), puis en écrivant que $\boldsymbol{r}_j = \boldsymbol{p}_j - \beta_{j-1}\boldsymbol{p}_{j-1}$ d'après (4.6c), on trouve

$$\begin{aligned} \boldsymbol{r}_{k+1}^\top \boldsymbol{r}_j &= (\boldsymbol{r}_k - \alpha_k \mathrm{A}\boldsymbol{p}_k)^\top \boldsymbol{r}_j \\ &= \boldsymbol{r}_j^\top \boldsymbol{r}_k - \alpha_k \boldsymbol{r}_j^\top \mathrm{A}\boldsymbol{p}_k \\ &= \boldsymbol{r}_j^\top \boldsymbol{r}_k - \alpha_k \boldsymbol{p}_j^\top \mathrm{A}\boldsymbol{p}_k + \alpha_k \beta_{j-1} \boldsymbol{p}_{j-1}^\top \mathrm{A}\boldsymbol{p}_k \end{aligned}$$

D'après l'hypothèse de récurrence on a forcément $\boldsymbol{p}_{j-1}^\top \mathrm{A}\boldsymbol{p}_k = 0$. Par ailleurs si $j < k$, encore d'après l'hypothèse de récurrence, on a $\boldsymbol{r}_j^\top \boldsymbol{r}_k = \boldsymbol{p}_j^\top \mathrm{A}\boldsymbol{p}_k = 0$. On vient donc de montrer que $\boldsymbol{r}_{k+1}^\top \boldsymbol{r}_j = 0$ pour $j < k$. Dans le cas $j = k$, on trouve $\boldsymbol{r}_{k+1}^\top \boldsymbol{r}_k = |\boldsymbol{r}_k|_2^2 - \alpha_k |\boldsymbol{p}|_{\mathrm{A}}^2$, et alors l'expression de α_k donnée par (4.6) nous permet de conclure que $\boldsymbol{r}_{k+1}^\top \boldsymbol{r}_k = 0$. En conclusion on a établit que

$$\boldsymbol{r}_{k+1}^\top \boldsymbol{r}_j = 0 \quad \forall j = 0 \ldots k. \tag{4.8}$$

Démontrons maintenant qu'on a également $\boldsymbol{p}_{k+1}^\top \mathrm{A}\boldsymbol{p}_j = 0$ pour $j = 0 \ldots k$. D'après (4.6c), et puisque $\mathrm{A}\boldsymbol{p}_j = (\boldsymbol{r}_j - \boldsymbol{r}_{j+1})/\alpha_j$ selon (4.6b), on obtient

$$\begin{aligned} \boldsymbol{p}_{k+1}^\top \mathrm{A}\boldsymbol{p}_j &= (\boldsymbol{r}_{k+1} + \beta_k \boldsymbol{p}_k)^\top \mathrm{A}\boldsymbol{p}_j \\ &= \boldsymbol{r}_{k+1}^\top \mathrm{A}\boldsymbol{p}_j + \beta_k \boldsymbol{p}_k^\top \mathrm{A}\boldsymbol{p}_j \\ &= \boldsymbol{r}_{k+1}^\top (\boldsymbol{r}_j - \boldsymbol{r}_{j+1})/\alpha_j + \beta_k \boldsymbol{p}_k^\top \mathrm{A}\boldsymbol{p}_j \end{aligned} \tag{4.9}$$

Le premier terme du membre de droite ci-dessus est nul puisque dans tous les cas on a $\boldsymbol{r}_{k+1}^\top \boldsymbol{r}_j = 0$. Par ailleurs si $j < k$, on a d'une part $\boldsymbol{r}_{k+1}^\top \boldsymbol{r}_{j+1} = 0$ d'après (4.8), et d'autre part $\boldsymbol{p}_k^\top \mathrm{A}\boldsymbol{p}_j = 0$ d'après l'hypothèse de récurrence. On en tire $\boldsymbol{p}_{k+1}^\top \mathrm{A}\boldsymbol{p}_j = 0$ si $j < k$.

Enfin, dans le cas où $j = k$, en combinant (4.9) avec les expressions de α_k, β_k provenant de (4.6), on déduit

$$\begin{aligned} \boldsymbol{p}_{k+1}^\top \mathrm{A}\boldsymbol{p}_k &= -\frac{1}{\alpha_k}|\boldsymbol{r}_{k+1}|_2^2 + \beta_k |\boldsymbol{p}_k|_\mathrm{A}^2 \\ &= -\frac{|\boldsymbol{p}_k|_\mathrm{A}^2}{|\boldsymbol{r}_k|_2^2}\, |\boldsymbol{r}_{k+1}|_2^2 + \frac{|\boldsymbol{r}_{k+1}|_2^2}{|\boldsymbol{r}_k|_2^2}\, |\boldsymbol{p}_k|_\mathrm{A}^2 = 0. \end{aligned}$$

Ceci clôt notre raisonnement par récurrence, et démontre que (4.7) est vrai pour tout $k \geq 0$. Notons que $\dim(\mathfrak{K}_{k+1}) \leq 1 + \dim(\mathfrak{K}_k)$ et $\dim(\mathfrak{K}_1) = 1$, de sorte que $\dim(\mathfrak{K}_k) \leq k$. Comme la première propriété de (4.7) montre que $\boldsymbol{r}_j, j = 0 \dots k-1$ est une famille de k vecteurs linéairement indépendants appartenant à $\mathfrak{K}_k$, c'est donc une base de $\mathfrak{K}_k$. On voit donc que la propriété $\boldsymbol{r}_k^\top \boldsymbol{r}_j = 0 \forall j = 0 \dots k-1$ revient à écrire $\boldsymbol{r}_k^\top \boldsymbol{y} = 0$ pour tout $\boldsymbol{y} \in \mathfrak{K}_k$. C'est précisément ce qui restait à démontrer pour conclure définitivement la preuve. □

Proposition 4.5.1.
Soit $\mathrm{A} \in \mathbb{R}^{n\times n}$ *une matrice symétrique définie positive. On note* $|\boldsymbol{x}|_\mathrm{A}^2 := \boldsymbol{x}^\top \mathrm{A}\boldsymbol{x}$. *Étant donné* $\boldsymbol{x}^{(0)}, \boldsymbol{b} \in \mathbb{R}^n$, *si* $(\boldsymbol{x}^{(k)})_{k\geq 0}$ *est la suite construite par la méthode du gradient conjugué (4.6),*

$$|\boldsymbol{x}^{(k)} - \boldsymbol{x}_\star|_\mathrm{A} \leq 2\Big(\frac{\sqrt{\mathrm{cond}_2(\mathrm{A})} - 1}{\sqrt{\mathrm{cond}_2(\mathrm{A})} + 1}\Big)^k |\boldsymbol{x}^{(0)} - \boldsymbol{x}_\star|_\mathrm{A} \qquad \forall k \geq 0.$$

Démo :

On a vu que la méthode du gradient conjugué est une méthode de projection orthogonale. A chaque itération, $\boldsymbol{x}^{(k)}$ peut donc être caractérisé comme la solution d'un problème de minimisation. En notant $\boldsymbol{d}_0 = \boldsymbol{x}^{(0)} - \boldsymbol{x}_\star$ on a

$$\begin{aligned} |\boldsymbol{x}^{(k)} - \boldsymbol{x}_\star|_\mathrm{A}^2 &= \min_{\boldsymbol{x}\in\boldsymbol{x}^{(0)}+\mathfrak{K}_k} |\boldsymbol{x} - \boldsymbol{x}_\star|_\mathrm{A}^2 \\ &= \min_{\mathrm{Q}\in\mathbb{P}_{k-1}[X]} |\boldsymbol{x}^{(0)} - \boldsymbol{x}_\star - \mathrm{Q}(\mathrm{A})\boldsymbol{r}_0|_\mathrm{A}^2 \\ &= \min_{\mathrm{Q}\in\mathbb{P}_{k-1}[X]} |(\boldsymbol{x}^{(0)} - \boldsymbol{x}_\star) - \mathrm{Q}(\mathrm{A})\mathrm{A}(\boldsymbol{x}^{(0)} - \boldsymbol{x}_\star)|_\mathrm{A}^2 \\ &= \min_{\mathrm{Q}\in\mathbb{P}_{k-1}[X]} |(\mathrm{Q}(\mathrm{A})\mathrm{A} + \mathrm{Id})\boldsymbol{d}_0|_\mathrm{A}^2 \\ &= \min_{\substack{\mathrm{Q}\in\mathbb{P}_k[X] \\ \mathrm{Q}(0)=1}} |\mathrm{Q}(\mathrm{A})\boldsymbol{d}_0|_\mathrm{A}^2 \end{aligned}$$

Remarquons maintenant que $|\mathrm{Q}(\mathrm{A})\boldsymbol{d}_0|_\mathrm{A}^2 = (\mathrm{Q}(\mathrm{A})\boldsymbol{d}_0)^\top \mathrm{A}(\mathrm{Q}(\mathrm{A})\boldsymbol{d}_0) = \boldsymbol{d}_0^\top \mathrm{Q}(\mathrm{A})\mathrm{A}\mathrm{Q}(\mathrm{A})\boldsymbol{d}_0$. Comme A est symétrique définie positive, elle est diagonalisable en base orthonormée. Il existe donc une matrice diagonale $\Lambda = \mathrm{diag}_{j=\dots n}(\lambda_j)$ et une matrice orthogonale $\mathrm{O} \in \mathbb{R}^{n\times n}, \mathrm{O}^\top = \mathrm{O}^{-1}$, telle que $\mathrm{A} = \mathrm{O}^{-1}\Lambda\mathrm{O}$. On en déduit $|\mathrm{Q}(\mathrm{A})\boldsymbol{d}_0|_\mathrm{A}^2 = (\mathrm{O}\boldsymbol{d}_0)^\top \mathrm{Q}(\Lambda)\Lambda\mathrm{Q}(\Lambda)(\mathrm{O}\boldsymbol{d}_0)$. Notons $\boldsymbol{y} = \mathrm{O}\boldsymbol{d}_0 = (y_j)_{j=1\dots n}$. On a donc

$$\begin{aligned} |\mathrm{Q}(\mathrm{A})\boldsymbol{d}_0|_\mathrm{A}^2 &= \textstyle\sum_{j=1}^n |\mathrm{Q}(\lambda_j)|^2 \lambda_j |y_j|^2 \\ &\leq (\max_{j=1\dots n} |\mathrm{Q}(\lambda_j)|)^2 \sum_{j=1}^n \lambda_j |y_j|^2 \\ &\leq (\max_{\lambda\in\mathfrak{S}(\mathrm{A})} |\mathrm{Q}(\lambda)|)^2 |\boldsymbol{d}_0|_\mathrm{A}^2 \end{aligned}$$

En conclusion, on vient d'obtenir

$$|\boldsymbol{x}^{(k)} - \boldsymbol{x}_\star|_{\mathrm{A}} = |\boldsymbol{x}^{(0)} - \boldsymbol{x}_\star|_{\mathrm{A}} \min_{\substack{\mathrm{Q}\in\mathbb{P}_k[X]\\ \mathrm{Q}(0)=1}} \max_{\lambda\in\mathfrak{S}(\mathrm{A})} |\mathrm{Q}(\lambda)|$$

Il reste donc à estimer le dernier terme danx le membre de droite ci-dessus. Il s'agit d'un problème classique d'interpolation. Notons $\lambda_1 = \max\mathfrak{S}(\mathrm{A})$ la plus grande valeur propre de A, et $\lambda_n = \min\mathfrak{S}(\mathrm{A})$ la plus petite. Une borne est fournie par la solution d'un problème classique de théorie d'interpolation (voir par exemple Appendice B, p.639 de [1])

$$\min_{\substack{\mathrm{Q}\in\mathbb{P}_k[X]\\ \mathrm{Q}(0)=1}} \max_{\lambda\in\mathfrak{S}(\mathrm{A})} |\mathrm{Q}(\lambda)| \leq \min_{\substack{\mathrm{Q}\in\mathbb{P}_k[X]\\ \mathrm{Q}(0)=1}} \max_{\lambda_n\leq\lambda\leq\lambda_1} |\mathrm{Q}(\lambda)| = \frac{1}{C_k\left(\frac{\lambda_1+\lambda_n}{\lambda_1-\lambda_n}\right)}. \tag{4.10}$$

Ici $C_k(t)$ désigne le k-ième polynôme de Chebyschev. Il s'agit de l'unique polynôme de degré k vérifiant l'identité $C_k(\cos\theta) = \cos(k\theta)$ pour $\theta\in\mathbb{R}$. Une expression explicite est donnée par la formule

$$C_k(t) = \frac{1}{2}[(t+\sqrt{t^2-1})^k + (t+\sqrt{t^2-1})^{-k}] \geq \frac{1}{2}(t+\sqrt{t^2-1})^k \tag{4.11}$$

Posons $\eta = \lambda_n/(\lambda_1-\lambda_n)$ de sorte que $(\lambda_1+\lambda_n)/(\lambda_1-\lambda_n) = 1+2\eta$. On peut alors encore re-écrire $t+\sqrt{t^2-1}$ comme

$$\begin{aligned} t+\sqrt{t^2-1} &= 1+2\eta+\sqrt{4\eta+4\eta^2} = \eta+(1+\eta)+2\sqrt{\eta}\sqrt{1+\eta} \\ &= (\sqrt{\eta}+\sqrt{1+\eta})^2 = \frac{(\sqrt{\lambda_1}+\sqrt{\lambda_n})^2}{\lambda_1-\lambda_n} = \frac{\sqrt{\lambda_1}+\sqrt{\lambda_n}}{\sqrt{\lambda_1}-\sqrt{\lambda_n}} \\ &= \frac{\sqrt{\lambda_1/\lambda_n}+1}{\sqrt{\lambda_1/\lambda_n}-1} = \frac{\sqrt{\mathrm{cond}_2(\mathrm{A})}+1}{\sqrt{\mathrm{cond}_2(\mathrm{A})}-1} \end{aligned} \tag{4.12}$$

En combinant (4.11) avec (4.12), et en injectant cette borne dans (4.10), on arrive au résultat annoncé. □

4.5.2 Gradient conjugué préconditionné

Rappelons que dans le cas où l'on cherche à résoudre l'équation $\mathrm{A}\boldsymbol{x} = \boldsymbol{b}$, préconditionner consiste à multiplier l'équation à gauche par une matrice M^{-1}, ce qui conduit à $\mathrm{M}^{-1}\mathrm{A}\boldsymbol{x} = \mathrm{M}^{-1}\boldsymbol{b}$, en choisissant M de manière à ce que $\mathrm{cond}_2(\mathrm{M}^{-1}\mathrm{A})$ soit plus proche de 1 que $\mathrm{cond}_2(\mathrm{A})$.

On pourrait penser qu'un tel procédé pose problème pour appliquer la méthode du gradient conjugué car même si M et A sont symétrique définie positives (SDP), il n'y pas de raison que $\mathrm{M}^{-1}\mathrm{A}$. Cependant la propéiété "être symétrique définie positive" doit s'entendre relativement à un produit scalaire, et ici il s'agit implicitement du produit scalaire canonique (celui associé à $|\ \ |_2$). Si l'on pose $(\boldsymbol{x},\boldsymbol{y})_{\mathrm{M}} := \boldsymbol{y}^\top\mathrm{M}\boldsymbol{x}$ alors on a

$$\begin{aligned} (\mathrm{M}^{-1}\mathrm{A}\boldsymbol{x},\boldsymbol{y})_{\mathrm{M}} &= \boldsymbol{y}^\top\mathrm{M}\mathrm{M}^{-1}\mathrm{A}\boldsymbol{x} = \boldsymbol{y}^\top\mathrm{A}\boldsymbol{x} = (\mathrm{A}\boldsymbol{y})^\top\boldsymbol{x} \\ &= (\mathrm{M}\mathrm{M}^{-1}\mathrm{A}\boldsymbol{y})^\top\boldsymbol{x} = (\mathrm{M}^{-1}\mathrm{A}\boldsymbol{y})^\top\mathrm{M}^\top\boldsymbol{x} \\ &= (\mathrm{M}^{-1}\mathrm{A}\boldsymbol{y})^\top\mathrm{M}\boldsymbol{x} = (\boldsymbol{x},\mathrm{M}^{-1}\mathrm{A}\boldsymbol{y})_{\mathrm{M}} \end{aligned}$$

En résumé on a $(\mathrm{M}^{-1}\mathrm{A}\boldsymbol{x}, \boldsymbol{y})_{\mathrm{M}} = (\boldsymbol{x}, \mathrm{M}^{-1}\mathrm{A}\boldsymbol{y})_{\mathrm{M}}$ c'est-à-dire que $\mathrm{M}^{-1}\mathrm{A}$ est symétrique définie positive vi-à-vis du produit scalaire $(\ ,\ )_{\mathrm{M}}$. En se basant sur cette remarque, on peut proposer une version préconditionnée de l'algorithme du gradient conjugué qui s'écrit alors

$$\boldsymbol{r}_0 = \mathrm{A}\boldsymbol{x}^{(0)} - \boldsymbol{b}, \quad \boldsymbol{z}_0 = \mathrm{M}^{-1}\boldsymbol{r}_0, \quad \boldsymbol{p}_0 = \boldsymbol{z}_0$$

$$\begin{aligned}
&\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} - \alpha_k \boldsymbol{p}_k\\
&\boldsymbol{r}_{k+1} = \boldsymbol{r}_k - \alpha_k \mathrm{A}\boldsymbol{p}_k\\
&\boldsymbol{z}_{k+1} = \mathrm{M}^{-1}\boldsymbol{r}_{k+1}\\
&\boldsymbol{p}_{k+1} = \boldsymbol{z}_{k+1} + \beta_k \boldsymbol{p}_k\\
&\text{avec}\\
&\alpha_k = \frac{\boldsymbol{r}_k^\top \boldsymbol{z}_k}{\boldsymbol{p}_k^\top \mathrm{A}\boldsymbol{p}_k} \quad \text{et} \quad \beta_k = \frac{\boldsymbol{r}_{k+1}^\top \boldsymbol{z}_{k+1}}{\boldsymbol{r}_k^\top \boldsymbol{z}_k}
\end{aligned}$$

4.5.3 Generalised Minimal Residual (GMRes)

La méthode GMRes est une méthode de projection oblique sur l'espace de Krylov. Elle construit ue suite de vecteurs $\boldsymbol{x}^{(k)}$ à partir d'un $\boldsymbol{x}^{(0)} \in \mathbb{R}^n$ et de $\boldsymbol{r}_0 = \mathrm{A}\boldsymbol{x}^{(0)} - \boldsymbol{b}$ suivant l'algorithme

$$\begin{cases} \boldsymbol{x}^{(k)} \in \boldsymbol{x}^{(0)} + \mathfrak{K}_k \\ (\mathrm{A}\boldsymbol{y})^\top(\mathrm{A}\boldsymbol{x}^{(k)} - \boldsymbol{b}) = 0 \quad \forall \boldsymbol{y} \in \mathfrak{K}_k. \end{cases} \tag{4.13}$$

D'après le lemme 4.3.1 on a

$$|\mathrm{A}\boldsymbol{x}^{(k)} - \boldsymbol{b}|_2 = \min_{\boldsymbol{x} \in \boldsymbol{x}^{(0)} + \mathfrak{K}_k} |\mathrm{A}\boldsymbol{x} - \boldsymbol{b}|_2$$

On sait que $\boldsymbol{r}_0, \mathrm{A}\boldsymbol{r}_0, \mathrm{A}^2\boldsymbol{r}_0, \ldots, \mathrm{A}^{k-1}\boldsymbol{r}_0$ est une base de $\mathfrak{K}_k$. On considère la base $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ de veteurs orthonormaux construits de la manière suivante

$$\begin{cases} \boldsymbol{v}_1 = \boldsymbol{r}_0/|\boldsymbol{r}_0|_2 \\ \boldsymbol{w}_{j+1} = \mathrm{A}\boldsymbol{v}_j - \sum_{p=1}^{j} \boldsymbol{v}_p \boldsymbol{v}_p^\top \mathrm{A}\boldsymbol{v}_j \\ \boldsymbol{v}_{j+1} = \boldsymbol{w}_{j+1}/|\boldsymbol{w}_{j+1}|_2 \end{cases}$$

Il s'agit du procédé d'orthonormalisation de Gram-Schmidt. Par construction $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ sont orthonormaux et forment une base de $\mathfrak{K}_k$. Notons alors

$$\begin{aligned}
&\mathrm{Q}_m := [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m]\\
&\boldsymbol{h}_j = (\mathrm{Q}_j^\top \mathrm{A}\boldsymbol{v}_j, |\boldsymbol{w}_{j+1}|_2, 0, \ldots, 0)^\top \in \mathbb{R}^{k+1}\\
&\mathrm{H}_k := [\boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_k] \in \mathbb{R}^{(k+1)\times k}
\end{aligned}$$

Selon ces définitions, si on écrit le vecteur $\boldsymbol{h}_j \in \mathbb{R}^{k+1}$ composante par composante $\boldsymbol{h}_j = (h_{p,j})_{p=1\ldots k+1}$, on a $h_{p,j} = \boldsymbol{v}_p^\top \mathrm{A}\boldsymbol{v}_j$ pour $p \leq j$, ainsi que $h_{j+1,j} = |\boldsymbol{w}_{j+1}|_2$ et $h_{p,j} = 0$ pour $p \geq j+2$. Soulignons également que H_k n'est pas une matrice carré, mais bien une matrice rectangulaire avec une ligne de plus que de colonnes. En re-écrivant l'algorithme de Gram-Schmidt ci-dessus avec ces nouvelles notations, pour $j \leq k$ on obtient donc :

$$\mathrm{A}\boldsymbol{v}_j = h_{j+1,j}\boldsymbol{v}_{j+1} + \sum_{p=1}^{j} h_{p,j}\boldsymbol{v}_p = \sum_{p=1}^{k+1} h_{p,j}\boldsymbol{v}_p = \mathrm{Q}_{k+1}\boldsymbol{h}_j$$

$$\begin{aligned}
\mathrm{A}\mathrm{Q}_k &= \mathrm{A}[\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k] = [\mathrm{A}\boldsymbol{v}_1, \ldots, \mathrm{A}\boldsymbol{v}_k]\\
&= [\mathrm{Q}_{k+1}\boldsymbol{h}_1, \ldots, \mathrm{Q}_{k+1}\boldsymbol{h}_k] = \mathrm{Q}_{k+1}[\boldsymbol{h}_1, \ldots, \boldsymbol{h}_k] = \mathrm{Q}_{k+1}\mathrm{H}_k
\end{aligned}$$

En résumé on a donc $\mathrm{AQ}_k = \mathrm{Q}_{k+1}\mathrm{H}_k$. Remarquons à présent que par définition de l'espace de Krylov, on recherche $\boldsymbol{x}^{(k)}$ sous la forme $\boldsymbol{x}^{(k)} = \boldsymbol{x}^{(0)} - \mathrm{Q}_k\boldsymbol{y}_k$ pour un certain $\boldsymbol{y}_k \in \mathbb{R}^k$, et que $\boldsymbol{r}_0 = |\boldsymbol{r}_0|_2\boldsymbol{v}_1 = \beta\mathrm{Q}_{k+1}\boldsymbol{e}_1$ où l'on a posé $\beta = |\mathrm{A}\boldsymbol{x}^{(0)} - \boldsymbol{b}|_2 = |\boldsymbol{r}_0|_2$ et $\boldsymbol{e}_1 = (1, 0 \dots, 0)^\top \in \mathbb{R}^{k+1}$ est le premier vecteur de la base canonique de $\mathbb{R}^{k+1}$.

Par ailleurs, observons que $|\mathrm{Q}_{k+1}\boldsymbol{y}|_2 = |\boldsymbol{y}|_2$ pour tout $\boldsymbol{y} \in \mathbb{R}^{k+1}$ car les colonnes Q_{k+1} sont orthonormées. En reprenant les notations que nous venons d'introduire, on a alors

$$\begin{aligned}
|\mathrm{A}\boldsymbol{x}^{(k)} - \boldsymbol{b}|_2 &= |\mathrm{A}\boldsymbol{x}^{(0)} - \boldsymbol{b} - \mathrm{AQ}_k\boldsymbol{y}_k|_2 = |\boldsymbol{r}_0 - \mathrm{AQ}_k\boldsymbol{y}_k|_2 \\
&= |\beta\mathrm{Q}_{k+1}\boldsymbol{e}_1 - \mathrm{Q}_{k+1}\mathrm{H}_k\boldsymbol{y}_k|_2 = |\mathrm{Q}_{k+1}(\beta\boldsymbol{e}_1 - \mathrm{H}_k\boldsymbol{y}_k)|_2 \\
&= |\beta\boldsymbol{e}_1 - \mathrm{H}_k\boldsymbol{y}_k|_2
\end{aligned}$$

En reprenant la caractérisation de $\boldsymbol{x}^{(k)}$ comme solution d'un problème de minimisation dans l'espace de Krylov, le calcul ci-dessus nous ramène à un problème de minimisation dans $\mathbb{R}^k$ qui s'écrit

$$\begin{aligned}
& |\mathrm{A}\boldsymbol{x}^{(k)} - \boldsymbol{b}|_2 = \min_{\boldsymbol{x} \in \boldsymbol{x}^{(0)} + \mathfrak{K}_k} |\mathrm{A}\boldsymbol{x} - \boldsymbol{b}|_2 \\
\iff \quad & |\mathrm{H}_k\boldsymbol{y}_k - \beta\boldsymbol{e}_1|_2 = \min_{\boldsymbol{y} \in \mathbb{R}^k} |\mathrm{H}_k\boldsymbol{y} - \beta\boldsymbol{e}_1|_2
\end{aligned}$$

Pour résumer, ceci nous amène finalement à considérer l'algorithme suivant pour le calcul de chaque $\boldsymbol{x}^{(k)}$.

$\boldsymbol{r}_0 = \mathrm{A}\boldsymbol{x}^{(0)} - \boldsymbol{b}$
$\beta = |\boldsymbol{r}_0|_2$
$\boldsymbol{v}_1 = \boldsymbol{r}_0/\beta$

FOR $j = 1 \dots k$
  $\boldsymbol{w}_j = \mathrm{A}\boldsymbol{v}_j$
  FOR $p = 1 \dots j$
    $h_{p,j} = \boldsymbol{v}_p^\top\boldsymbol{w}_j$
    $\boldsymbol{w}_j = \boldsymbol{w}_j - h_{p,j}\boldsymbol{v}_p$
  ENDFOR
  $h_{j+1,j} = |\boldsymbol{w}_j|_2$
  $\boldsymbol{v}_{j+1} = \boldsymbol{w}_j/h_{j+1,j}$
ENDFOR

$\mathrm{H}_k = (h_{p,j})_{p=1\dots k+1, j=1\dots k}$
Calculer $\boldsymbol{y}_k$ en minimisant $\boldsymbol{y} \mapsto |\mathrm{H}_k\boldsymbol{y} - \beta\boldsymbol{e}_1|_2$
$\boldsymbol{x}^{(k)} = \boldsymbol{x}^{(0)} - \mathrm{Q}_k\boldsymbol{y}_k$.

Bibliographie

[1] O. Axelsson. *Iterative Solution Methods.* Cambridge University Press, New York, NY, USA, 1994.

[2] T.H. Cormen, C.E. Leiserson, R.L. Rivest, and C. Stein. *Introduction to algorithms.* MIT Press, Cambridge, MA, third edition, 2009.

[3] N. M. Josuttis. *The C++ Standard Library : A Tutorial and Reference.* Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1999.