

Méthode et Pratiques Scientifiques

Nahid Ehmadi

Table des matières

1	Introduction	2
1.1	Principaux éléments	2
1.2	But du cours	2
2	Quel paradigme de programmation ?	3
2.1	Une approche par composants...	3
2.2	... eux-même liés par un graph	3
3	Unir et conquérir	3
4	Multiple Krylov Subspace	4
5	Résolution de problèmes de grande taille	4
6	Méthode de la puissance	5

Référence :

— *Introduction au Calcul Matriciel et Optimisations*, P. Ciarlet

1 Introduction

Technique de regroupement On cherche à partitionner les sommets V d'un graphe $G = (V, E)$ dans un ensemble de clusters $S_k \subseteq V$ tel que $\bigcup_{k=1}^p S_k = V$. La modularité est une mesure de la qualité d'un partitionnement des noeuds dans les communautés.

La plupart du temps, les solutions à ce problèmes sont NP-Complexe : on les approxime par à l'aide d'algorithmes, qui sont principalement basées sur du calcul de valeurs propres. Ce dernier prend 90% du temps total ! De plus, 90% du temps passé dans le solveur s'effectue dans des multiplications (sparse) matrice-vecteur .

1.1 Principaux éléments

- Méthode itératives pour problèmes de grande taille
- Méthodes hybride synchrone/asynchrone
- Méthode numériques d'algèbre linéaire pour le traitement de masses de données (simulation de phénomènes physiques, analyse des réseaux sociaux, etc)
- Méthode de compression des structures creuses
- Modèles de programmation de graphe de Tâches, PGAS
- Métriques de performances

Quelques rappels sur la méthode de Gauss

Méthode directe vs Itérative Une méthode directe est une solution dont la solution peut être calculée avec un nombre fini d'opérations arithmétiques élémentaires. A contrario, une méthode itératives est un procédé qui part d'une information initiale arbitraire et renvoie un résultat approché. Pour cela, on réinjecte le résultat en entrée de l'algorithme afin de raffiner la solution.

Exemple Pour trouver les valeurs propres d'une matrice, il faut trouver les racines du polynôme caractéristique, ce qui est impossible dès que le polynôme dépasse le degré 4 : on n'utilise donc jamais une résolution directe.

La plupart du temps, les matrices sont creuses : on n'utilise donc pas les mêmes méthodes, plus particulièrement pas de méthodes directes car celle-ci ne sont pas adaptées.

Une des problématiques de nos jours est la consommation électrique. La nouvelle frontière du *calcul exascale* impose des consommations énormes. La machine la plus puissante au 11/16 consommait plus de 15MW !

1.2 But du cours

Le principe de ce cours est la mise au point d'un paradigme de calcul pour la programmation à grande échelle et par la suite en déduire l'introduction de méthodes numériques modernes et "intelligentes" comme celles Krylov hybrides smart-tunées.

2 Quel paradigme de programmation ?

Certains chercheurs pensent directement à OpenMP/MPI lorsque l'on parle d'informatique parallèle. C'est *faux* : il ne s'agit que d'*outils de programmation*, qui ne sont donc pas représentatifs de la théorie derrière.

Par exemple, une multiplication matrice-vecteur puis réduction par addition du vecteur obtenu, il faut faire communiquer ensemble tous les processus : impossible. On rappelle que le principal coût énergétique est le déplacement des données !

On cherche donc à :

- Minimiser le mouvement des données (amoindrir au maximum le coût des communications)
- Consommation énergétique
- Parallélisme multi-niveau
- Ordonnancement multi-niveau
- Équilibrage de charge multi-niveau
- Tolérance aux pannes

2.1 Une approche par composants...

Il faudrait un modèle de programmation laissant la possibilité de définir des blocs autonomes (cf GLCS) appelés *composants* réutilisables dans différents contextes. Par exemple, un composant pour le produit matrice-vecteur creux peut être réutilisé dans un solveur calculant les valeurs propres. Si le travail est bien fait, on peut même réutiliser un composant séquentiel dans un programme parallèle.

2.2 ... eux-même liés par un graph

On peut définir un problème comme un ensemble de composants liés entre eux par un graph. L'application est donc un ensemble de composants *de très gros grains*, car chaque composant peut être lui-même un graph de composants.

Un projet de recherche à ce sujet est YML (attention, encore expérimental).

Les principales caractéristiques du modèle de programmation recherché sont alors :

- Privilégier grandement les communication asynchrones sur les communications synchrones
- Prendre en compte l'hétérogénéité des unités de calcul
- Prendre en compte la tolérance aux fautes
- Encourager l'équilibrage de tâche
- Introduire l'auto-tuning (on choisit automatiquement des paramètres pendant l'exécution avec assistance de l'utilisateur final)

3 Unir et conquérir

On prend pour exemple la résolution d'un système linéaire $Ax = b$, où $A \in \mathcal{M}_n(\mathbb{C})$ et $b, x \in \mathbb{C}$ sont creux. On cherche quelques *paires de Ritz* $\Lambda_k = (\lambda_1, \dots, \lambda_k)$ et $U_k = (u_1, \dots, u_k)$ tel que $\forall i \leq k, Au_i = \lambda_i u_i$.

Ce problème est au cœur de l'algorithme de recherches de données de Google.

Problème : Les algorithmes classiques ne sont pas adaptables aux grandes tailles (par exemple, la méthode QR pour trouver les valeurs propres a une précision dépendant de la taille de la matrice). On va donc chercher à réduire notre problème dans un sous-espace de taille moindre, chercher la solution dans ce sous-espace puis effectuer une transformation inverse afin d'obtenir une approximation des solutions recherchés. Le plus souvent, la précision n'est pas suffisante, et on va définir un sous-espace de meilleur qualité pour réitérer le processus.

Pour notre exemple, on va projeter notre matrice dans un sous-espace de Krylov défini par la base $(v, Av, \dots, A^{m-1}v)$. On résout le problème de valeur propre dans ce sous-espace ; puis on réitère le processus avec le v trouvé en résolution du système de moins grande taille.

Attention, il ne s'agit pas vraiment d'une méthode itérative, mais plus d'une méthode "à redémarrage" (projection-résolution-retour à l'espace de départ).

La méthode peut être redémarrée explicitement (ERAM/Saad), c'est-à-dire en mettant à jour explicitement v puis recommençant le problème. Par contre, cette méthode n'est pas toujours stable. Un redémarrage implicite peut aussi être effectué (IRAM/Sorrensens) ; mais repose sur des bases mathématiques plus conséquentes.

Ces méthodes ont par contre un inconvénient : si les valeurs propres sont trop approchées les unes des autres, l'algo peut les "confondre".

La méthode *Unite and conquer* consiste à utiliser plusieurs méthodes.

4 Multiple Krylov Subspace

On va non pas combiner un ensemble de méthode, mais combiner différentes instances de la même méthode. Par exemple, on va projeter-résoudre-retourner, puis on partage les résultats intermédiaires, et on choisit les meilleurs vecteurs retournés par les instances. On peut également choisir des tailles de sous-espaces différentes de manière à recouvrir les temps des communications (asynchrones).

La sélection d'une bonne valeur pour la dimension du sous-espace peut être faite par auto-tuning.

5 Résolution de problèmes de grande taille

Soit $A \in \mathbb{C}^n$ une matrice creuse. Le problème est de calculer un petit nombre s de valeurs propres λ_i et ses vecteurs propres (normalisés) correspondants v_i de $A : Av_i = \lambda_i v_i$ pour $i \in \llbracket 1, s \rrbracket$. On suppose les valeurs propres triées par ordre décroissant de module. On note $\text{lin}(s)$ pour désigner tous les vecteurs colinéaires à s .

Définition 1 (Suite de Krylov). Soit q un vecteur non nul. La suite $(A^n q)^{n \in \mathbb{N}}$ est appelée suite de Krylov.

Définition 2 (Sous-espace de Krylov). Le sous espace de Krylov de taille m est le sous-espace engendré par les m premiers vecteurs de la suite de Krylov si ces vecteurs sont linéairement indépendants.

Théorème 1.

$$A^m q \xrightarrow{m \rightarrow +\infty} \text{lin}(v_1)$$

La méthode de la puissance est aussi appelée méthode des itérations du quotient de Rayleigh. Il s'agit de calculer $A^m q$ pour $m = 1, 2, \dots$ jusqu'à convergence.

Méthode des itérations simultanées (IS) Soit $S = (q_1 \ \dots \ q_k)$ où q_1, \dots, q_k sont des vecteurs linéairement indépendants. On calcule S, AS, A^2S, \dots . On peut montrer que la suite $A^m S \xrightarrow{m \rightarrow +\infty} \text{lin}(v_1, v_2, \dots, v_k)$.

Algorithme d'Arnoldi On définit les sous-espaces de Kyrlov $K_m(a, q_i) = \lim(q_1, Aq_1, \dots, A^{m-1}q_1)$. On calcul en fait une base orthonormée.

```

1  $q_1 = \frac{q}{||q||}$ 
2 for  $k=1$  à  $m-1$  do
3    $w = Aq_k$ 
4   for  $j=1$  à  $k$  do
5      $h_{j,k} = \langle w, q_j \rangle$ 
6      $w = w - h_{j,k} \cdot q_j$ 
7    $h_{k+1,k} = ||w||_2$ 
8    $q_{k+1} = \frac{w}{h_{k+1,k}}$ 

```

L'algorithme fournit alors H_{m,m_1} et Q_m tels que $AQ_{m-1} = Q_m \cdot H_{m,m-1}$, que l'on peut réécrire en $AQ_m = Q_{m-1} \cdot H_{m-1,m-1} + q_m \cdot h_{m,m-1} \cdot e_{m-1}^T$ où le second terme est le *résidu* (et e_{m-1} le $m-1$ ème vecteur de la base canonique).

Remarques

- Si $h_{m,m} = 0 \Rightarrow Q_{m-1}^T A Q_{m-1} = H_{m-1,m-1}$ et les valeurs propres de $H_{m-1,m-1}$ sont celles de A .
- Si $m-1 = n \Rightarrow Q - n$ est une base orthonormal de \mathbb{C}^n et les valeurs propres de $H_{n,n}$ sont celles de A .
- Soit $H = H_{m-1,m-1}$ et $Hy = \lambda y$ et $w = Q_{m-1}y$ alors l'équation devient

$$Av = \lambda v + h_{m,m-1} q_m e_{m-1}^T y$$

Soit

$$||Av - \lambda v|| = |h_{m,m-1}| \cdot |y_{m-1}| \cdot \underbrace{||q_m||_2}_{=0}$$

Le résidu est alors $||Av - \lambda v||_2 = |h_{m,m-1}| \cdot |y_{m-1}|$

6 Méthode de la puissance

Soit A une matrice carrée de taille n et de valeurs propres $\lambda_1, \dots, \lambda_n$ ses valeurs propres associés aux vecteurs v_1, \dots, v_n triées par modules décroissant. Soit q un vecteur non orthogonal à v_1 . Alors l'algorithme de la méthode de la puissance est

```

1  $q_0 = \frac{q}{||q||}$ ;
2 for  $i = 0$  à  $n$  do
3    $q_{i+1} = A \frac{q_i}{||Aq_i||}$ ;

```

Théorème 2. Si $q_0(1) \neq 0$ (première composante de q_0), alors $q_n \xrightarrow{n \rightarrow \infty} v_1$ à un facteur multiplicatif près.

Démonstration. On note $\gamma_n = ||Aq_n||^{-1}$

$$\begin{aligned}
q_0 &= \sum_{i=1}^n \alpha_i v_i \\
q_1 &= Aq_0 \cdot \gamma_0 \\
&= \gamma_0 A \sum_{i=1}^n \alpha_i v_i = \gamma_0 \sum_{i=1}^n \alpha_i \underbrace{Av_i}_{\lambda_i v_i}
\end{aligned}$$

De même,

$$\begin{aligned} q_2 &= \gamma_1 \gamma_0 A \sum_{i=1}^n \alpha_i \lambda_i v_i \\ &= \gamma_1 \gamma_0 \sum_{i=1}^n \alpha_i \lambda_i^2 v_i \end{aligned}$$

D'où

$$\begin{aligned} q_m &= \underbrace{\left(\prod_{i=0}^{m-1} \gamma_i \right)}_{=\zeta_{m-1}} \cdot \sum_{i=1}^n \alpha_i^m v_i \\ &= \lambda_1^m \sum_{i=1}^n \alpha_i \cdot \zeta_{m-1} \cdot \underbrace{\left(\frac{\lambda_i}{\lambda_1} \right)^m}_{\xrightarrow[n \rightarrow +\infty]{0} 0 \text{ quand } i \neq 1} v_i \end{aligned}$$

Or $\zeta_{m-1} = \prod_{i=0}^{m-1} \gamma_i = \prod_{i=0}^{m-1} \|Aq_i\|^{-1} \leq \prod_{i=0}^{m-1} \|\lambda_1 \cdot q_i\|^{-1} \leq |\lambda_1|^{-m}$
Il vient

$$\lim_{m \rightarrow +\infty} q_m = \lim_{m \rightarrow +\infty} \lambda_1^m \zeta_{m-1} \alpha_1 v_1$$

□

Remarque La vitesse de convergence est régie par $\frac{\lambda_2}{\lambda_1}$.

Propriété 1. Soit $\mu \in Sp(A)$. Alors la matrice $(A - \mu I)^{-1}$ possède les mêmes vecteurs propres que A et ses valeurs propres sont $(\lambda - \mu)^{-1}$ où λ est une valeur propre de A .

Corollaire 1. Si μ est proche d'une valeur propre λ_j de A , Alors $(\lambda_j - \mu)^{-1}$ sera une valeur propre beaucoup plus grande que $(\lambda_i - \mu)^{-1}$ pour $i \neq j$. On peut appliquer la méthode de la puissance à la matrice $A - \lambda I$ et accélérer ainsi sa vitesse de convergence.

Méthode de déflation

Soit μ un nombre complexe. On cherche à déterminer la valeur propre de A la plus proche de μ . Pour cela, on définit une suite de vecteur $q_{n+1} = (A - \mu I)^{-1} q_n$. On peut démontrer que $\langle q_{n+1}; q_n \rangle \rightarrow \lambda$ et $\frac{q_n}{\|q_n\|} \rightarrow v$ avec λ la valeur propre la plus proche de μ et v sont vecteur propre associé.

Supposons λ_1 et v_1 connus, nous recherchons λ_2 et v_2 .

Cas symétrique Posons $B = A - \lambda_1 \left(\frac{v_1 \cdot v_1^T}{\|v_1\|^2} \right)$. Les valeurs propres de B sont $0, \lambda_2, \dots, \lambda_n$. Il suffit donc d'appliquer la méthode des puissances à B .

Cas non symétrique B n'as pas de raison d'avoir les $0, \lambda_2, \dots, \lambda_n$ comme valeurs propres (car les vecteurs propres de A ne sont pas forcément orthogonaux). On utilise dans ce cas la matrice transposée de A (qui a les même valeurs propres que A). Soit $A^T w_1 = \lambda_1 w_1$. Alors $B = A - \left(\frac{v_1 w_1^T}{\|w_1^T v_1\|} \right)$.

```
clock_t start, end;  
start = clock();  
//program here  
end = clock();  
printf("%d\n", (end-start)/CLOCKS_PER_SEC);
```