

Chapitre 4 - Matlab et analyse numérique

IV

Dans ce chapitre nous abordons les notions centrales de l'analyse numérique. En effet, beaucoup de méthodes numériques conduisent à la résolution d'un système d'équations (linéaire ou non linéaire). Quelques méthodes de résolution des équations sont détaillées ici, ainsi que des méthodes d'intégration et manipulation des polynômes.

1. Fonctions "numériques"

Les fonctions numériques de Matlab généralisent les fonctions numériques usuelles, avec une différence cependant, elles sont vectorisées. C'est à dire qu'elles s'appliquent aussi bien à des nombres qu'à des tableaux. Lorsqu'une de ces fonctions a pour argument un tableau, la fonction est appliquée à chacun de ces éléments. Le résultat obtenu est donc un tableau du même format que le tableau donné comme argument.

2. Polynômes

Sur Matlab, les vecteurs ligne sont utilisés pour représenter les polynômes, en effet, un polynôme étant déterminé par ses coefficients, il suffit de stocker ceux-ci. Le coefficient du monôme de degré 0 dans la première coordonnée du vecteur et le coefficient du monôme de plus haut degré dans la dernière. Par exemple pour représenter le polynôme $P(x) = 2x^2 - 3x + 1$ on utilise le vecteur $[2 \ -3 \ 1]$.

2.1. Manipuler les polynômes

La fonction la plus utilisée sur les polynômes est évidemment celle qui permet d'évaluer la valeur d'un polynôme P pour une valeur x donnée. La fonction `polyval` prend comme arguments un polynôme p (sous forme d'un vecteur) et un nombre x , et fournit comme résultat la valeur $p(x)$.

Exemple : Évaluer un polynôme

Dans cet exemple nous allons évaluer la valeur du polynôme $P(x) = 2x^2 - 3x + 1$ pour $x = 2$.

```
1 >> P = [2 -3 1];
2 >> polyval(P, 2)
3
4 ans =
5
6      3
```

Complément

D'autres fonctions qui sont assez répandues dans Matlab pour manipuler les polynômes sont listées dans le tableau suivant :

| Fonction | Opération à effectuer |
|-------------------------------|---|
| <code>poly(r)</code> | Créer un polynôme à partir de ces racines (dans le vecteur r passé comme argument). |
| <code>polyfit(x, y, n)</code> | Retourner un polynôme p de degré n donnant la meilleure approximation (en méthode des moindres carrés) pour les points définis par les vecteurs x et y . |
| <code>roots(p)</code> | Retourner les racines du polynôme p passé comme argument. |
| <code>conv(p, q)</code> | Effectuer une multiplication de p par q . |
| <code>deconv(p, q)</code> | Effectue une division euclidienne du polynôme p par q . Il est aussi possible de demander deux sorties à la fonction <code>deconv</code> pour calculer aussi le résidu de la division polynomiale. En d'autres termes, la commande <code>[u, v] = deconv(p, q)</code> donne deux polynômes u et v tel que : $p = \text{conv}(u, q) + v$. |
| <code>polyint(p)</code> | Retourne l'intégrale du polynôme p (en utilisant la valeur 0 comme constante d'intégration). Il est aussi possible de spécifier la constante d'intégration k en la passant comme deuxième argument de la fonction : <code>(polyint(p, k))</code> . |
| <code>polyder(p)</code> | Retourner le polynôme dérivé de p . |

Tableau 4 : Les fonctions les plus utilisées sur les polynômes

Exemple

Nous montrons dans cet exemple quelques commandes effectuant des manipulations sur les polynômes

```
1 >> P = [1 -1];      % P(x) = x - 1.
2 >> Q = [-1 2];     % Q(x) = x - 1.
3 >> W = conv(P, Q)  % W(x) = P(x) * Q(x).
4
5 W =
6
7     -1     3    -2
8
9 >> r = roots(W)    % Calculer les racines de W
10
11 r =
12
13     2
14     1
15
16 >> v = poly(r)    % créer un polynôme à partir de ces racines
17
18 v =
19
20     1    -3     2
```

```

21
22 >> U = deconv(W, P) % U(x) = W(x) / P(x)
23
24 U =
25
26     -1     2
27
28 >> A = polyder(V) % Calculer le polynôme A(x) = V'(x)
29
30 A =
31
32     2    -3
33
34 >> polyint(A) % Calculer l'intégral de A(x)
35
36 ans =
37
38     1    -3     0
39
40 >> polyint(A, 2) % Calculer l'intégral de A(x) avec la constante d'intégration
41     2
42 ans =
43
44     1    -3     2

```

3. Calcul matriciel

Matlab est spécialement conçu pour manipuler des matrices. Il reconnaît et manipule les variables matricielles à coefficients réels ou complexes, denses ou creuses. Nous avons déjà montré dans le premier chapitre comment initialiser, effectuer des calculs et appliquer des fonctions sur des matrices. Nous montrons dans cette section comment effectuer des calculs matriciels plus avancés en utilisant Matlab.

3.1. Les matrices creuses

On appelle matrice creuse une matrice comportant une forte proportion de coefficients nuls. L'intérêt de telles matrices résulte principalement de la réduction de la place mémoire (on ne stocke pas les zéros) et aussi de la réduction du nombre d'opérations (on n'effectuera pas les opérations portant sur les zéros). Par défaut dans MATLAB une matrice est considérée comme pleine, c'est-à-dire que tous ses coefficients sont mémorisés. Si M est une matrice, la commande `sparse(M)` permet d'obtenir la même matrice mais stockée sous la forme creuse. Si l'on a une matrice stockée sous la forme creuse, on peut obtenir la même matrice, stockée sous la forme ordinaire par la commande `full`.

Exemple : Avantage des matrices creuses

Dans cet exemple nous allons utiliser la commande `whos` pour avoir des informations sur l'espace mémoire occupé par chaque variable, puis nous comparons l'espace mémoire occupé par des matrices creuses stockées sous différents formats.

```

1 >> A = diag(1:10)
2
3 A =
4
5     1     0     0     0     0     0     0     0     0     0

```

```

6     0     2     0     0     0     0     0     0     0     0
7     0     0     3     0     0     0     0     0     0     0
8     0     0     0     4     0     0     0     0     0     0
9     0     0     0     0     5     0     0     0     0     0
10    0     0     0     0     0     6     0     0     0     0
11    0     0     0     0     0     0     7     0     0     0
12    0     0     0     0     0     0     0     8     0     0
13    0     0     0     0     0     0     0     0     9     0
14    0     0     0     0     0     0     0     0     0    10
15
16 >> B = sparse(A)
17
18 B =
19
20    (1,1)      1
21    (2,2)      2
22    (3,3)      3
23    (4,4)      4
24    (5,5)      5
25    (6,6)      6
26    (7,7)      7
27    (8,8)      8
28    (9,9)      9
29    (10,10)    10
30
31 >> infoA = whos('A'); % Récupérer des informations sur la variable A
32 >> infoA.bytes      % Afficher le nombre de bytes (octets) occupés par A
33
34 ans =
35
36    800
37
38 >> infoB = whos('B'); % Récupérer des informations sur la variable B
39 >> infoB.bytes      % Afficher le nombre de bytes (octets) occupés par B
40
41 ans =
42
43    164

```

3.2. Valeurs et vecteurs propres d'une matrice

Soit A une matrice carrée d'ordre n . On dit que x est un vecteur propre de la matrice A pour la valeur δ si $Ax = \delta x$. Pour trouver les valeurs propres de la matrice A , il faut résoudre le système $Ax = \lambda x \iff (A - \lambda I_n)x = 0$. Avec I_n une matrice identité de taille n (matrice carrée de taille n dont tous les éléments de la diagonale valent "1" et tous les autres éléments valent "0").

À l'exclusion de la solution triviale $x = 0$, le système ci-dessus admet des solutions si le déterminant de $(A - \lambda I_n)$ vaut 0.

 **Complément : Calculer le vecteur propre associé à chaque valeur propre**

Un vecteur x est associé comme vecteur propre à la valeur λ s'il vérifie la relation $(A - \lambda I_n)x = 0$.



Complément : Calculer les vecteurs et les valeurs propres sur Matlab

Pour calculer les valeurs propres d'une matrice carrée A nous utilisons la fonction `eig(A)` qui retourne un vecteur colonne contenant toutes les valeurs propres de A . Il est aussi possible de calculer les valeurs et les vecteurs propres de A en demandant deux sorties à la fonction `eig([vec, val] = eig(A))`. la deuxième sortie de la fonction `eig` dans ce cas donne une matrice diagonale contenant les valeurs propre de A , alors que la première sortie donne une matrice telle que chaque colonne est un vecteur propres de A .

Plus formellement, la commande `[vec, val] = eig(A)` assure que $A * \text{vec} = \text{vec} * \text{val}$.



Exemple : Utilisation de la fonction eig

Nous montrons dans cet exemple comment utiliser la fonction `eig` de Matlab pour calculer les vecteurs et valeurs propres d'une matrice

```
1 >> A = [5 -3; 6 -4];
2 >> [vec, val] = eig(A)
3
4 vec =
5
6     0.7071     0.4472
7     0.7071     0.8944
8
9
10 val =
11
12     2     0
13     0    -1
14
15 >> A * vec
16
17 ans =
18
19     1.4142    -0.4472
20     1.4142    -0.8944
21
22 >> vec * val
23
24 ans =
25
26     1.4142    -0.4472
27     1.4142    -0.8944
```

3.3. Méthodes itératives pour la résolution des systèmes linéaires

Un système d'équations linéaires est un ensemble d'équations portant sur les mêmes inconnues. Nous considérons ici que les systèmes contenant autant d'équations que de variables. En général, un système de m équations linéaires à n inconnues peut être écrit sous la forme suivante :

Implémentation Matlab

Nous montrons ici l'implémentation Matlab d'une fonction permettant de résoudre un système d'équations linéaires donné sous sa forme matricielle. Pour simplifier l'implémentation, nous allons décomposer la matrice A en deux sous-matrices, la matrice diagonale D contenant les éléments diagonaux de A , et $R = A - D$. Dans ce cas, le système précédent est équivalent à :

$$x^{(k+1)} = D^{-1}(b - Rx^{(k)}) \text{ pour } k \geq 0 \quad (4)$$

```
1 % Fonction - jacobi.m
2 function x = jacobi(A, b, X0)
3 D = diag(diag(A)); % D prend les éléments diagonaux de A
4 invD = D^(-1); % nous calculons D^-1
5 R = A - D; % R contient les éléments non diagonaux de A
6 Xk = X0;
7 while norm(b - A * Xk) > eps % condition d'arrêt
8     Xkplus1 = invD * (b - R * Xk); % calculer Xkplus1 en fonction de k
9     Xk = Xkplus1;
10 end;
11 x = Xk;
```

3.3.2. Méthode de Gauss-Seidel

La méthode de Gauss-Seidel est très similaire à celle de Jacobi, la seule différence est qu'elle propose d'introduire au fur et à mesure dans le calcul, sans attendre l'itération suivante, les composantes de $x^{(k+1)}$ qui viennent d'être calculées. Le schéma numérique de la méthode de Jacobi est le suivant :

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \times \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij} x_j^{(k)} \right) \text{ pour } k \geq 0 \quad (5)$$

L'algorithme de Gauss-Seidel est en général plus rapide que l'algorithme de Jacobi, et donc préférable. Par contre l'algorithme de Jacobi est plus adapté aux environnements parallèles.

4. Fonction d'une variable

Nous montrons dans ce chapitre les méthodes numériques pour le calcul d'intégrale, et pour la résolution des équations non linéaires. On ne considère dans cette section que les fonctions à une seule variable.

4.1. Calcul d'intégrale

Très souvent le calcul explicite de l'intégrale, d'une fonction f continue sur un intervalle $[a, b]$, définie par : $I(f) = \int_a^b f(x) dx$ peut être très difficile à atteindre. Par conséquent, on fait appel à des méthodes numériques afin de calculer une approximation de $I(f)$. Nous présentons ici deux méthodes numériques pour le calcul d'intégrale : méthode du point milieu et méthode du trapèze.

4.1.1. Méthode du point milieu

Soit f une fonction continue sur l'intervalle $[a, b]$. L'idée de cette méthode est de subdiviser l'intervalle $[a, b]$ en n sous intervalles $I_k = [x_k, x_{k+1}]$ tel que $x_1 = a$ et $x_{n+1} = b$. Le schéma numérique de la méthode du point milieu est le suivant :

$$I(f) = \sum_{k=1}^n (x_{k+1} - x_k) \times f(\bar{x}_k) \quad \text{avec} \quad \bar{x}_k = \frac{x_{k+1} + x_k}{2} \quad (6)$$

En effet, $I(f)$ représente l'aire comprise entre la courbe $y = f(x)$ et l'axe des abscisses entre les droites $x = a$ et $x = b$. Alors que comme le montre la figure ci-dessous, $(x_{k+1} - x_k) \times f(\bar{x}_k)$ calcule l'aire entre la droite $y = f(\bar{x}_k)$ et l'axe des abscisses sur l'intervalle I_k . Plus les intervalles I_k sont réduits (et par conséquent nombreux), plus l'approximation du calcul d'intégrale est précise.

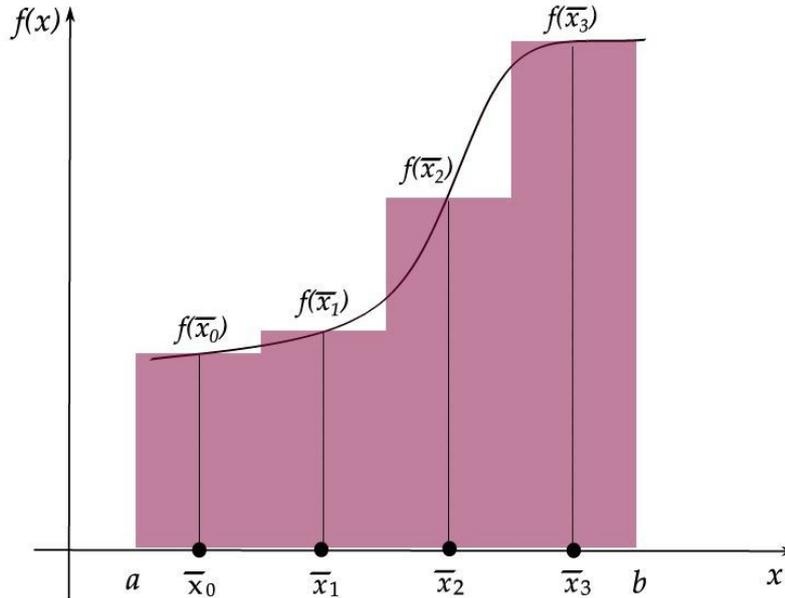


Figure 2 : Représentation de la méthode du point milieu

Si tous les intervalles I_k sont équidistants, le schéma numérique peut être simplifié comme suit :

$$I(f) = \Delta x \times \sum_{k=1}^n f(\bar{x}_k) \quad \text{avec} \quad \Delta x = \frac{b-a}{n} \quad (7)$$

Implémentation Matlab

Nous montrons ici l'implémentation Matlab d'une fonction permettant de calculer l'intégrale d'une autre fonction (passée comme paramètre) en utilisant la méthode du point milieu. Pour que notre calcul d'intégrale soit assez dynamique, il doit prendre en charge toutes les fonctions qui ont un seul paramètre x . Pour cela, la fonction pour laquelle nous allons calculer l'intégrale est passée aussi comme paramètre, ainsi que les bornes de l'intervalle a et b (nous divisons l'intervalle $[a, b]$ par la suite en 100 sous intervalles équidistants). La fonction est donnée sous forme d'une chaîne de caractères, puis elle est convertie à une fonction en appelant `str2func`.

```
1 % Fonction - pointMilieu.m
2 function res = pointMilieu(strFunc, a, b)
3 integFunc = str2func(strFunc);
4 Xk = linspace(a, b, 101); % Générer tous les Xk distribués
   uniformement sur l'intervalle [a, b]
5 dx = (b - a) / 100; % Calculer Δx
6 mp = (Xk(2:end) + Xk(1:end-1)) ./ 2; % Calculer les centres des sous intervalles
7 res = sum(integFunc(mp)) * dx; % Calculer l'integrale
```

4.1.2. Méthode du trapèze

La méthode du trapèze consiste aussi à subdiviser l'intervalle $[a, b]$ en n sous intervalles I_k . Cette méthode est basée sur l'interpolation linéaire de chaque intervalle. En d'autres termes, sur chaque intervalle $[x_k, x_{k+1}]$ est substitué par la droite joignant les points $(x_k, f(x_k))$ et $(x_{k+1}, f(x_{k+1}))$. Le schéma numérique de la méthode du trapèze est le suivant :

$$I(f) = \sum_{k=1}^n \left(\frac{f(x_k) + f(x_{k+1})}{2} \times (x_{k+1} - x_k) \right) \quad (8)$$

La figure ci-dessous montre que $((f(x_{k+1}) + f(x_k))/2) \times (x_{k+1} - x_k)$ calcule la surface de la trapézoïdale (d'où le nom trapèze) délimitée par la droite $((x_k, f(x_k)), (x_{k+1}, f(x_{k+1})))$ et l'axe des abscisses sur l'intervalle I_k . Il est aussi possible d'augmenter la précision de la méthode du trapèze en réduisant la taille des intervalles I_k .

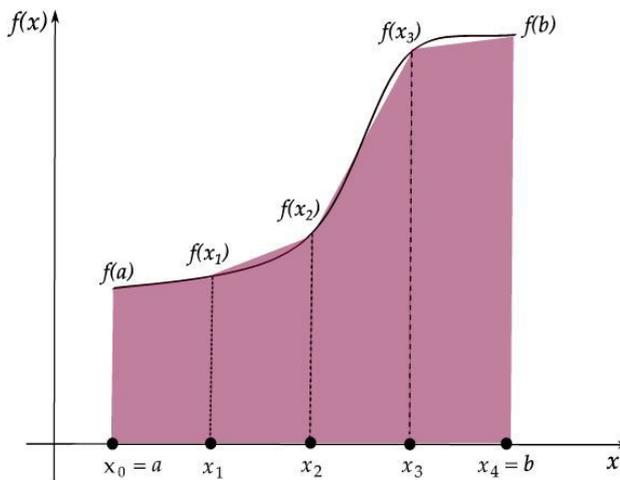


Figure 3 : Représentation de la méthode du trapèze

Si tous les intervalles I_k sont équidistants, le schéma numérique peut être simplifié comme suit :

$$\begin{aligned} I(f) &= \Delta x \times \sum_{k=1}^n \frac{f(x_k) + f(x_{k+1})}{2} \quad \text{avec} \quad \Delta x = \frac{b-a}{n} \\ &= \frac{\Delta x}{2} \times \sum_{k=1}^n (f(x_k) + f(x_{k+1})) \\ &= \frac{\Delta x}{2} \times \left(\sum_{k=1}^n f(x_k) + \sum_{k=1}^n f(x_{k+1}) \right) \\ &= \frac{\Delta x}{2} \times \left(\sum_{k=1}^n f(x_k) + \sum_{k=2}^{n+1} f(x_k) \right) \\ &= \frac{\Delta x}{2} \times \left(f(x_1) + \sum_{k=2}^n f(x_k) + \sum_{k=2}^n f(x_k) + f(x_{n+1}) \right) \\ &= \frac{\Delta x}{2} \times \left(f(x_1) + f(x_{n+1}) + 2 \sum_{k=2}^n f(x_k) \right) \\ &= \frac{\Delta x}{2} \times \left(f(a) + f(b) + 2 \sum_{k=2}^n f(x_k) \right) \\ &= \Delta x \times \left(\frac{f(a) + f(b)}{2} + \sum_{k=2}^n f(x_k) \right) \end{aligned} \quad (9)$$

La méthode du trapèze est prédéfinie dans l'environnement Matlab par la fonction `trapz`. Elle est utilisée selon la syntaxe suivante : `trapz(x, y)`. L'argument `x` de la fonction `trapz` contient les x_k qui définissent les subdivisions de l'intervalle sur lequel on va calculer l'intégral, alors que l'argument `y` contient les $f(x_k)$.

4.2. Résolution des équations non-linéaires

Plusieurs méthodes numériques permettent de calculer une approximation des zéros d'une fonction f . Deux méthodes sont illustrées dans cette section : méthode du point fixe et de la dichotomie.

4.2.1. Méthode du point fixe

Le principe de la méthode du point fixe est que de transformer l'équation $f(x) = 0$ à une équation $g(x) = x$ (généralement $g(x) = f(x) + x$). Donc, résoudre l'équation $f(x) = 0$ revient à trouver α tel que $g(\alpha) = \alpha$. Dans le cas où α existe, il est qualifié de point fixe de la fonction $g(x)$ qui est appelée *fonction d'itération*. Le schéma numérique de la méthode du point fixe est le suivant :

$$x^{(k+1)} = g(x^{(k)}) \quad \text{pour } k > 0 \quad (10)$$

Notez que $x^{(0)}$ est généralement choisi d'une façon aléatoire sur un intervalle $[a, b]$.

Implémentation Matlab

Nous donnons ici une fonction qui prend les trois paramètres d'entrées suivants :

- f : Le nom d'une fonction sous forme d'une chaîne de caractères ;
- a et b : les bornes de l'intervalle sur lequel on doit chercher la solution.

Et qui retourne comme résultat α tel que $f(\alpha) = \alpha$ en utilisant la méthode du point fixe.

```
1 % Fonction - pointFixe.m
2 function x = pointFixe(strFunc, a, b)
3 f = str2func(strFunc);
4 g = @(x) f(x) + x;           % Définir la fonction g(x) = f(x) + x
5 x = rand() * (b - a) + a;    % Choisir un point du départ aléatoire
6 while x ~= g(x)             % Répéter x = g(x) jusqu'à trouver le point fixe
7     x = g(x);
8 end
```

⚠ Attention : Test d'égalité

Il est important de noter que les ordinateurs n'utilisent pas explicitement les nombres réels. Ils utilisent plutôt les nombres flottants, qui sont un sous-ensemble discret de l'ensemble des nombres réels. Donc il est possible que α (tel que $g(\alpha) = \alpha$) n'appartient pas à ce sous-ensemble. Pour cela il est recommandé d'éviter les tests explicites d'égalité (ou d'inégalité), et de les remplacer par des tests qui sont assez légers pour être salifiables (et en même temps assez exigeants pour que l'algorithme reste correct).

Dans l'exemple du script précédent, il est recommandé de remplacer la condition `x ~= g(x)` par `(x - g(x)) / x < eps`.

4.2.2. Méthode de dichotomie

Le principe de cette méthode est basé sur le théorème des valeurs intermédiaires.

◆ *Rappel : Théorème des valeurs intermédiaires*

Ce théorème implique que si une fonction f est continue sur un intervalle $[a, b]$, et que $f(a)$ et $f(b)$ ont différents signes ($f(a) * f(b) < 0$) alors il existe un nombre α qui appartient à $[a, b]$ tel que $f(\alpha) = 0$.

La méthode de dichotomie suit les étapes suivantes :

- Calculer $c = (a + b) / 2$.
- Si $f(c) = 0$ retourner c comme résultat.
- Sinon, si $\text{signe}(c) = \text{signe}(a)$ (a et c ont le même signe), alors continuer le travail récursivement sur l'intervalle $[c, b]$, sinon continuer sur l'intervalle $[a, c]$.

