Comment programmer de façon efficace et quelques particularités de Matlab

6-601-09 Simulation Monte Carlo

Geneviève Gauthier

HEC Montréal

Introduction
La précision
Le temps de calcul
Matlab et les boucles
La gestion de la mémoire vive
Projet d'envergure
Les bases de la programmation en Matlab
Les fonctions
L'optimisation
Retour à la précision

Introduction

Peu importe le langage utilisé, il y a deux composantes importantes liées à l'utilisation de méthodes numériques :

1. La précision
2. Le temps de calcul.

Introduction
La précision
Le temps de calcul
Matlab et les boucles
La gestion de la mémoire vive
Projet d'envergure
Les bases de la programmation en Matlab
Les fonctions
L'optimisation
Retour à la précision

La précision

Les erreurs d'arrondissement

1. La précision dépend, entre autres, de la façon dont un logiciel traite les nombres réels.
2. Matlab entrepose tous les nombres selon le "IEEE floating-point standard".
3. Les nombres en point-flottant ont une précision d'environ 16 chiffres significatifs et sont compris entre $\pm 10^{-308}$.

Par conséquent, il peut y avoir des erreurs d'arrondissement.

Introduction
La précision
Le temps de calcul
Matlab et les boucles
La gestion de la mémoire vive
Projet d'envergure
Les bases de la programmation en Matlab
Les fonctions
L'optimisation
Retour à la précision

La précision

Erreurs d'approximation

La précision dépend de certaines approximations. Par exemple,

1. $\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!} \cong \sum_{k=0}^{K} \frac{x^k}{k!}$
2. La fonction de répartition de la loi normale (nous y reviendrons).

Quelques règles élémentaires permettront d'éviter certaines erreurs.

1. Éviter de soustraire des nombres du même ordre de grandeur.
2. Il vaut mieux sommer les petits nombres en premier puis les ajouter aux grands.

Introduction
La précision
Le temps de calcul
Matlab et les boucles
La gestion de la mémoire vive
Projet d'envergure
Les bases de la programmation en Matlab
Les fonctions
L'optimisation
Retour à la précision

Le temps de calcul

Fonctions Matlab

Voici quelques fonctions de Matlab qui vous permettront de mesurer l'efficacité de certaines portions de votre code.

1. clock
2. cputime
3. tic, toc

Voir la description de la rubrique d'aide de Matlab.

Introduction
La précision
Le temps de calcul
Matlab et les boucles
La gestion de la mémoire vive
Projet d'envergure
Les bases de la programmation en Matlab
Les fonctions
L'optimisation
Retour à la précision

Les opérateurs de base I

Certaines opérations arithmétiques sont beaucoup plus rapides que les autres. Le tableau suivant classe les opérations des plus rapides aux plus lentes.

Opérateur		Matlab
Additions et soustractions	$a + b$, $a - b$	$a + b$, $a - b$
Multiplications	$a \times b$	$a * b$
Divisions	$a \div b$	a / b
Exposants (puissance entière)	a^n	$x\hat{}n$
Exposants (général)	a^b	$x\hat{}b$

Les opérateurs de base II

Introduction
La précision
Le temps de calcul
Matlab et les boucles
La gestion de la mémoire vive
Projet d'envergure
Les bases de la programmation en Matlab
Les fonctions
L'optimisation
Retour à la précision

Voici quelques exemples courants qui pourront accélérer votre code :

Plus lent	Plus rapide
$2 * x$	$x + x$
$x/5$	$0.2 * x$
$x\hat{}3$	$x * x * x$
$a * (c + d) + b * (c + d)$	$(a + b) * (c + d)$
$a * x\hat{}2 + b * x + c$	$c + x * (b + a * x)$ (méthode d'Homer)
$a\hat{}0.5$	$sqrt(a)$
$a/b/c$	$a/(b * c)$

Évidemment, c'est lorsque cela est répété un très grand nombre de fois que la différence se fait sentir.

La méthode d'Homer pour les polynômes

Considérons le polynôme

$$f(x) = \sum_{i=0}^{n} a_i x^i \quad \text{(Methode 1)}$$

$$f(x) = a_0 + a_1 x + \sum_{i=2}^{n} a_i \left(\prod_{j=1}^{i} x \right) \quad \text{(Methode 2)}$$

$$f(x) = a_0 + x\left(a_1 + x\left(a_2 + \ldots\right)\right) \quad \text{(Methode 3)}$$

	Additions	Multiplications	Exponentiations
Méthode 1	n	n	$n-1$
Méthode 2	n	$n\left(n+1\right)/2$	0
Méthode 3	n	n	0

Introduction
La précision
Le temps de calcul
Matlab et les boucles
La gestion de la mémoire vive
Projet d'envergure
Les bases de la programmation en Matlab
Les fonctions
L'optimisation
Retour à la précision

Les opérateurs plus complexes

Les fonctions logarithme et exponentielle sont plus exigeantes en temps de calcul. Il faut utiliser les relations

Plus lent	Plus rapide
$\exp(x)\exp(y)$	$\exp(x+y)$
$\log(a)+\log(b)$	$\log(a*b)$
$\log(a)-\log(b)$	$\log(a/b)$
$\log(a\hat{}n)$	$n*\log a$

Introduction
La précision
Le temps de calcul
Matlab et les boucles
La gestion de la mémoire vive
Projet d'envergure
Les bases de la programmation en Matlab
Les fonctions
L'optimisation
Retour à la précision

Matlab et les boucles I

Matlab est un langage très efficace lorsqu'il s'agit d'opérations matricielles. Il est très lent lorsqu'il s'agit de boucles. Il faut donc les éviter le plus possible.

Loops.m

Mais si vous êtes *obligés* de créer des boucles, alors il faudrait suivre les règles suivantes :

1. Il faut que tout code invariant soit à l'extérieur de la boucle.
2. Lorsqu'il y a des boucles imbriquées, la plus "grosse" devrait être à l'intérieur de la plus "petite".
3. Les énoncés *if-then-else* devrait autant que possible être à l'extérieur de la boucle plutôt qu'à l'intérieur. De plus, la condition la plus fréquente devrait être vérifiée en premier.

Introduction
La précision
Le temps de calcul
Matlab et les boucles
La gestion de la mémoire vive
Projet d'envergure
Les bases de la programmation en Matlab
Les fonctions
L'optimisation
Retour à la précision

Matlab et les boucles II

4. Bien que Matlab permette de ne pas définir une matrice et de référer à ses éléments au fur et à mesure de sa construction, cela requiert significativement plus de temps de calcul.
5. Attention à la façon dont le logiciel gère les matrices. Elles sont généralement traitées comme des vecteurs et il devient préférable d'avoir plus de lignes que de colonnes.
6. Matlab est un langage matriciel et tout ce qui est possible d'être "vectoriser" devrait l'être.

Introduction
La précision
Le temps de calcul
Matlab et les boucles
La gestion de la mémoire vive
Projet d'envergure
Les bases de la programmation en Matlab
Les fonctions
L'optimisation
Retour à la précision

La gestion de la mémoire vive

Bien que les plus récents ordinateurs possèdent beaucoup de mémoire vive, une gestion intelligente de la mémoire pourrait grandement accélérer l'exécution de votre code.

1. Utiliser la commande "clear" pour effacer les variables dont vous ne vous servirez plus.
2. Ré-utiliser les mêmes variables temporaires.
3. Ne pas stocker inutilement de l'information.

Introduction
La précision
Le temps de calcul
Matlab et les boucles
La gestion de la mémoire vive
Projet d'envergure
Les bases de la programmation en Matlab
Les fonctions
L'optimisation
Retour à la précision

Projet d'envergure

Lorsque l'on s'attaque à un projet d'envergure, il faut écrire un pseudo-code décrivant en mots la structure du programme.

1. Les données à télécharger.
2. L'initialisation des variables.
3. Les principales fonctions, leurs entrées et leurs sorties.
4. Quels sont les résultats attendus et comment doivent-ils être présentés.
5. Il faudrait que la dépendance entre les sections du programme soit clairement identifiée.

Introduction
La précision
Le temps de calcul
Matlab et les boucles
La gestion de la mémoire vive
Projet d'envergure
Les bases de la programmation en Matlab
Les fonctions
L'optimisation
Retour à la précision

Projet d'envergure

Le programme doit ensuite être codé.

1. Utiliser les commentaires pour documenter votre programme (%) (et vos fonctions).
2. Le choix des noms de variables est important pour faciliter la compréhension du code.
3. L'indentation facilite la lecture du code.
4. Les codes qui seront utilisés à plusieurs reprises devraient être mis, autant que possible, sous forme de fonctions.

Introduction
La précision
Le temps de calcul
Matlab et les boucles
La gestion de la mémoire vive
Projet d'envergure
Les bases de la programmation en Matlab
Les fonctions
L'optimisation
Retour à la précision

Projet d'envergure

Vérifier que le programme calcule bien ce qu'il doit calculer en ...

1. ... exécutant le code sur des cas où vous connaissez la réponse (existence d'une solution analytique, reproduction de résultats publiés, etc.).
2. ... en vérifiant certaines propriétés théoriques
   - la parité put-call,
   - le comportement de la solution dans certains cas limites (lorsque la valeur du sous-jacent est nulle ou très grande),
   - si la solution est une fonction croissante d'un des paramètres, il faudrait le tester numériquement, etc.).

Améliorer l'efficacité de votre code en mesurant les temps d'exécution de certaines parties et en gérant le mieux possible les goulots d'étranglement.

Introduction
La précision
Le temps de calcul
Matlab et les boucles
La gestion de la mémoire vive
Projet d'envergure
Les bases de la programmation en Matlab
Les fonctions
L'optimisation
Retour à la précision

Les fonctions

- Il faut coder la fonction dans un fichier .m portant le même nom que la fonction. Dans le cas qui suit, le nom du fichier doit être funcname.m.

```
function [out1, out2, ...]  = funcname(in1,
in2, ...)
statement ...
```

- Les fonctions nargin et nargout permettent de contrôler les nombres d'intrants et d'extrants de la fonction.
- Les fonctions "anonymes" nous rendent un fier service...

L'optimisation

- Matlab possède plusieurs routines d'optimisation très efficaces.
- Il est important de bien connaître les techniques d'optimisation d'une routine avant de l'employer. Certaines sont adaptées à des problèmes particuliers.
- Ces fonctions d'optimisation retournent des variables qui nous renseignent sur la qualité de l'optimum obtenu:
  - Y a-t-il eu convergence?
  - Quel critère d'arrêt a été atteint?
  - etc.

Introduction
La précision
Le temps de calcul
Matlab et les boucles
La gestion de la mémoire vive
Projet d'envergure
Les bases de la programmation en Matlab
Les fonctions
L'optimisation
Retour à la précision

Estimation de la fonction de répartition de la loi normale

- La fonction de répartition est évaluée numériquement.
- Choisir une méthode qui a le niveau de précision approprié pour l'utilisation que l'on en fait. Méfiez-vous des ailes de la distribution.
- Certaines approximations sont beaucoup plus exigeantes en temps de calcul que d'autres.

Introduction
La précision
Le temps de calcul
Matlab et les boucles
La gestion de la mémoire vive
Projet d'envergure
Les bases de la programmation en Matlab
Les fonctions
L'optimisation
Retour à la précision