

**RAPPORT**  
**SUR**  
**L'ETUDE**  
**DES**  
**DONNEES FINANCIERES**  
**ET STATISTIQUES**  
**AL'AIDE**  
**DU LOGICIEL**  
**SCILAB**

**PAR :**

**MAROOF ASIM  
DAN BENTOLILA  
WISSAM ESSID**

**GROUPE 1 LM206 Lundi 10H45**

**INTRODUCTION:**

( Ce rapport est un compte rendu suite à la présentation du projet le lundi 18 décembre 2006)

Le but principal de ce projet consiste à récupérer des données financières sous forme d'un tableau et puis essayer d'étudier ces données mathématiquement grâce aux outils mathématiques à l'aide du logiciel Scilab. Concrètement, il s'agit par exemple de récupérer le taux d'échange entre l'Euro et le dollar au cours de l'année 2004 et de transformer ces données en forme de matrice dans Scilab pour mieux pouvoir étudier ces données. Grâce aux fonctions déjà prédéfinies dans le logiciel Scilab on pourra essayer d'ajuster la courbe obtenue par les données financières de départ. Ce travail est important à réaliser dans le domaine des statistiques car les données qu'on récupère au départ sont très lourdes et ne varient au premiers abord selon aucune loi mathématique connue ainsi on va essayer d'ajuster la courbe pour pouvoir l'approcher par une courbe mathématique aux lois connues (par exemple par une droite ou autre fonction polynomiale). Une fois la courbe transformée en une loi mathématique il est donc facile de prévoir l'évolution du marché et de comprendre son comportement tout au long de l'année.

La première tâche consiste donc à alléger ces données, par exemple pour le taux d'échange il faut transformer les données journalières de l'année en des données hebdomadaires. Ainsi on utilisera les fonctions déjà prédéfinies dans le langage Scilab pour les notions de moyenne et de variance dans le code Scilab. On pourra aussi utiliser la notion de moyenne mobile pour pouvoir comprendre la tendance de la courbe afin de prévoir l'évolution du taux d'échange dans les semaines à venir. Une fois ce travail réalisé il devient plus clair de voir le comportement général de la courbe c'est à dire de savoir si elle évolue comme une droite ou bien une autre fonction polynomiale. Ainsi grâce aux lois de la probabilité générale, on essaiera d'appliquer la méthode de la régression de droite à notre courbe du taux d'échange et essayer d'en faire une droite linéaire qui décrit approximativement le comportement de la courbe.

La dernière étape consiste à étudier la corrélation entre 2 actifs financiers, en effet une fois qu'on aura ajuster les paramètres des fonctions et qu'elles seront décrites par des courbes linéaires il sera facile de faire une corrélation entre les différents actifs financiers par exemple de savoir s'il y'a une dépendance entre le taux d'échange euro-dollar et dollar-yen.

## *I- Récupération de données*

La première utilité du logiciel Scilab est de pouvoir transformer des données économiques en objet mathématique c'est à dire en une matrice avec la quelle il est plus commode de travailler dans le langage Scilab.

Par exemple pour le taux d'échange entre l'euro et le dollar au cours de l'année 2004, on peut récupérer les données sous forme de tableau sur le site

<http://www.oanda.com>. Une fois ces données récupérées, on peut grâce à l'aide de la fonction prédéfinie « Read » transformer ce tableau en une matrice.

Exemple :

```
>data=read('dollar-euro-2004.txt',-1,1)'
```

nous créons une matrice "data" contenant toutes les valeurs du taux d'échange entre l'euro et le dollar au cours de l'année 2004, c'est une matrice de dimension 1\*366.

Grâce à la fonction plot2d, on peut obtenir la courbe des données et avoir un aperçu plus parlant de ces données financières.



## **TAUX D'ECHANGE EURO-DOLLAR 2004**

On emploie la notion d'écart type pour calculer une tendance centrale c'est à dire de savoir s'il y'a une dispersion autour de la moyenne ou pas. Dans scilab, pour calculer la moyenne on peut utiliser la fonction « mean » et pour l'écart type on s'aide la relation entre variance et l'écart type c'est à dire en utilisant le code :

$\text{sqrt}(\text{variance}(\text{data}))$ , ce qui nous donne en fonction de notre matrice data les valeurs suivantes :

```
-1->mean(data)
```

```
ans =  
0.8051022  
-1->sqrt(variance(data))  
ans =  
0.0264747
```

Ainsi on remarque que la tendance centrale de la courbe varie très sensiblement autour de très petite valeur et donc l'utilité de vouloir alléger ces données pour avoir quelques chose de très simple a manipuler.

## II- Calcul de la moyenne hebdomadaire et moyenne mobile

Comme on a vu dans le paragraphe précédent il est nécessaire d'effectuer des opérations sur les données de la courbe afin d'obtenir quelque chose de simple. Pour être précis et fiable dans l'étude le meilleur moyen reste de calculer la moyenne hebdomadaire. Ce procédé consiste à transformer la matrice 1\*366 en une matrice 7\*52 c'est à dire réparti par jour de semaine en fonction des semaines dans l'année et d'utiliser la fonction « mean » pour en faire une moyenne hebdomadaire, ainsi on passera de 366 données a 52 données ce qui sera plus agréable a manipuler.

Code : (pour créer une matrice 7\*52 en fonction de la matrice « data »)

```
tmp=[]  
l=7  
for i=1 :l  
tmp=[tmp;data(i:l:$-l+i)];  
end
```

puis il s'agit juste de faire la moyenne par ligne de la matrice « tmp » en utilisant la fonction « mean » :

```
z=mean((tmp),'r')
```

Ensuite grâce à la fonction plot2d et ses options, on peut faire un deuxième appel de plot2d de sorte qu'on ait la deuxième courbe c'est à dire plot2d de « z » dans la même fenêtre que le plot2d de « data ».

Premier appel :

```
Plot2d(a=length(data),data,style=5,frameflag=5,rect=[0,0.72,370,0.85])
```

Deuxième appel

```
B=[1 :7 :360]
```

```
Plot2d(b,z,style=2,frameflag=0) et on obtient:
```



On remarque que la courbe obtenue en bleu est assez proche de la courbe de départ et est moins lourde que la courbe de départ d'où l'intérêt principal de faire la moyenne hebdomadaire. Mais cette courbe n'est pas vraiment linéaire d'où la nécessité d'appliquer la méthode de régression linéaire pour simplifier davantage cette courbe. En revanche la courbe bleue nous permet de voir l'allure de la courbe qu'on obtiendra une fois la finition terminée.

Avec la même méthode on peut aussi obtenir une courbe décrivant la moyenne mensuelle mais celle ci ne sera pas très proche de la courbe réelle car l'écart est important entre 30 jours.

### **LA MOYENNE MOBILE**

En ce qui concerne la moyenne mobile son rôle est de montrer ou bien d'essayer de faire comprendre l'évolution du marché, c'est un indicateur important pour pouvoir comprendre les changements brusque des taux d'échange et aussi un bon moyen de comprendre les tendances et les variations du taux d'échange. En ce qui concerne notre projet il nous permettra de connaître la tendance globale de notre taux d'échange afin de faire une régression de droite la plus fiable possible. Ainsi on calculera la moyenne mobile sur 30 et 200 jours.

Pour calculer la moyenne mobile le principe reste toujours le même, il consiste à créer des matrices en fonction des jours selon le code suivant :

**h=**

```

l=30
for i = 1 : l
h=[h ;data(i :$-l+i)] ;
end

```

et ainsi on obtient les courbes suivantes :



Cette moyenne mobile sur 30 jours nous permet de voir que lorsqu'on va appliquer la méthode des régressions linéaire on pourra découper notre courbe en trois grande partie, la première allant du jour 0 à 150, la deuxième allant de 150 à 250 puis la troisième partie allant de 250 à 366.

### III- METHODE DES REGRESSION DE DROITE

La méthode des moindres carrées ou de régression de droite consiste à chercher une droite de type  $y=Ax+B$  en s'aidant des formules mathématiques utilisées en probabilités et statistique. On cherche A telle que soit égale à

$$A = \frac{\sum (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$

où  $\bar{x}$  est la moyenne des jours,  $\bar{y}$  la moyenne de la matrice pour les jours choisis,  $x_i$  est le jour et  $y_i$  est la valeur du taux d'échange correspondant au jour  $x_i$ .

et on peut retrouver B en considérant que la droite  $Ax+B$  passe par les points de moyenne( $\bar{x}$ ,  $\bar{y}$ ).

Le code qui permet d'obtenir la première droite sur la partie 1 à 150 s'écrit :

```
mx=(1 :150) (matrice de dim 150 correspondant aux 150 premiers jours)
ix=mean(mx) (correspond à  $\bar{x}$  )
sx=sum(((1 :150)-ix)^2) (correspond à  $\sum (x_i - \bar{x})^2$  )
my=data(1 :150) (matrice des données correspondant aux 150 premiers jours)
iy=mean(my)
sxy=sum(((1 :150)-ix).*(data(1 :150)-iy)) (correspondant à  $\sum (x_i - \bar{x}) \cdot (y_i - \bar{y})$  )

a=sxy/sx (correspond au coefficient A)
```

```
b=iy-a*ix ( on retrouve B par l'hypothèse que la droite linéaire passe par le point des moyennes)
```

Il faut appliquer le même code pour obtenir les 2 autres parties concernant les jours 150 à 250 et 250 à 366.

On obtient donc : (3 droites qui ne sont pas liés)



Une fois les droites obtenues, il est très facile maintenant d'étudier la courbe et de comprendre la variation générale.

Grâce à ces droites, on peut aussi étudier le résidu c'est à dire une matrice de dimension 1\*366 qui donne l'écart entre la droite théorique et les valeurs réelles.

Pour cela il faut calculer  $U_i = y'_i - y_i$

#### IV- Etude de Corrélation entre deux actifs financiers

Une fois les courbes simplifiées, il devient alors très facile d'étudier la corrélation entre deux actifs financiers. Par exemple on peut voir si le taux d'échange entre l'euro et le dollar a un effet sur la variation du prix de l'or. Il suffit pour cela de calculer le coefficient de corrélation qui est :

$$R = \frac{\sum (x_i - \bar{x}) \cdot (y_i - \bar{y})}{[\sum (x_i - \bar{x})^2]^{0,5} \cdot [\sum (y_i - \bar{y})^2]^{0,5}}$$

Désormais les valeurs  $y_i$  et  $x_i$  sont les valeurs des deux droites linéaires obtenues grâce à la régression des deux actifs financiers ce qui permet d'obtenir des valeurs très théoriques mais expliquant bien la corrélation entre les deux actifs financiers. En effet, si  $R$  est proche de  $-1$  on a que les deux actifs financiers varient en sens contraire et dépendent chacun de l'autre. Inversement si  $R=1$ , il y'a toujours dépendance mais varient dans le même sens et si  $R=0$ , on a absence de corrélation.



Sur Scilab, la fonction déjà prédéfinie « correl » n'est d'aucune utilité car on a besoin des matrices de fréquences ce qui n'est pas évident mais on peut utiliser la relation mathématique

$R = \text{Covariance}(x,y) / \text{Ecart type}(x) * \text{Ecart Type}(y)$   
avec x et y les actifs financiers.

#### ***V- REMARQUES sur scilab***

L'avantage de Scilab est qu'il possède énormément de fonctions prédéfinies sur les statistiques comme la fonction « mean », « variance », « correl ». Ce qui permet de faire des calculs assez compliqués et redondants avec une rapidité qui s'avère très efficace. De plus ça représente un gain de temps assez considérable vu que parfois le calcul des matrices peut être très long parfois.

Par contre, on s'aperçoit très vite qu'il a une sorte de mauvaise habitude dans la manipulation pratique. En statistique il y'a beaucoup de matrice à manipuler et le logiciel a tendance a réécrire toute la matrice qu'on lui affecte ce qui peut être très lourd et faire un travail très brouillon dans la manipulation.

Dans l'ensemble Scilab s'avère être un logiciel très utile surtout pour les étudiants car il est très simple à manipuler avec un vocabulaire ultra-simpliste, très différent de maple ou mathematica, ce qui permet de réaliser le travail de façon très efficace et surtout très rapidement.