Business Management
Gph-International Journal
www.gphjournal.org

10.5281/zenodo.7876780 Vol. 06 Issue 04 April - 2023 Paper Id: 826

DE LA DIFFÉRENTIATION FRACTIONNAIRE DES SÉRIES ÉCONOMIQUES

LAHCEN OULHAJ
PROFESSEUR DE SC. ÉCONOMIQUES

1. Introduction

Pour modéliser les séries économiques et afin d'éviter des corrélations fallacieuses, les chercheurs en économie procèdent souvent à des différentiations, première ou seconde selon l'ordre d'intégration des séries considérées, dans le but de les stationnariser. Généralement, ils commencent par administer des tests de racine unitaire (ADF, KPSS ou autre) aux séries. La plupart des séries économiques sont non stationnaires en niveau. Les séries sont alors différenciés en retranchant des valeurs y_t les valeurs y_{t-1}. Ils obtiennent ainsi les différences premières $\nabla y_t = y_t - y_{t-1}$. On teste la stationnarité de la nouvelle série. Si elle s'avère stationnaire, on travaille avec cette série des différences premières en l'introduisant dans un modèle de régression. Si la série des différences premières n'est pas stationnaire, on la différencie une nouvelle fois pour obtenir des différences de différences, ces différences secondes $\nabla^2 y_t = \nabla y_t - \nabla y_{t-1} = (y_t - y_{t-1}) - (y_{t-1} - y_{t-2}) = y_t - 2y_{t-1} + y_{t-2}$. Si ces différences secondes sont stationnaires, on travaille avec, sinon, on passe aux triples différences : $\nabla^3 y_t = \nabla.\nabla^2 y_t = \nabla^2 y_t - \nabla^2 y_{t-1} = (y_t - 2y_{t-1} + y_{t-2}) - (y_{t-1} - 2y_{t-2} + y_{t-3}) = y_t - 3y_{t-1} + 3y_{t-2} + y_{t-3}$
Lorsque la série est stationnaire en niveau, ce qui est rare, on dit qu'elle est intégrée d'ordre 0 et elle nécessite 0 différentiation pour l'intégrer. Lorsqu'elle nécessite une différentiation pour être stationnarisée, on dit qu'elle est intégrée d'ordre 1. Lorsqu'il faut une double différentiation pour être stationnarisée, la série est intégrée d'ordre 2, et ainsi de suite.
En vérité, il est rare qu'une série soit intégrée d'un ordre entier, 0, 1, 2,... Le plus souvent, l'ordre d'intégration des séries économiques est un nombre décimal ou "fractionnaire", compris entre 0 et 1 ou entre 1 et 2...
Ce papier portant sur la différentiation entière et fractionnaire des séries chronologiques présente, dans la section 2, la différentiation d'un ordre fractionnaire et, dans la section 3, l'application de cette différentiation aux séries temporelles économiques. Un algorithme de détermination de l'ordre fractionnaire d'intégration des séries avec son application à des séries macroéconomiques du Maroc est présenté dans la section 4. Dans la section 5, sont exposés les avantages de la différentiation fractionnaire des séries par rapport à la différentiation entière souvent adoptée par les chercheurs.

2. différentiation fractionnaire

Nous avons vu que les différences premières sont ainsi définies :

$\nabla = y_t - y_{t-1}$

LO-FRAC

Si on pose $y_{t-1} = Ly_t$ L étant l'opérateur de retard, alors

$\nabla y_t = y_t - Ly_t = (1-L)y_t$

On obtient ainsi pour les deuxièmes différences :

$\nabla^2 y_t = \nabla y_t - \nabla y_{t-1} = (1-L)y_t - (1-L)y_{t-1} = (1-L)(y_t - y_{t-1}) = (1-L)(1-L)y_t$

Nous avons vu ci-haut que $\nabla^2 y_t = y_t - 2y_{t-1} + y_{t-2} = (L^2 - 2L + 1)y_t$. En effet :

$(1-t)(1-t)y_t = (1-L)^2 y_t = (1-2L+L^2)y_t = \nabla^2 y_t$

On a de même $\nabla^3 y_t = (1-L)^3 y_t$

Les différences d'ordre n, $n \in \mathbb{N}$, de y_t sont donc obtenues comme ceci :

$\nabla^n y_t = (1-L)^n y_t = \sum_{k=0}^{n} \binom{n}{k} (-L)^k y_t$ (notation américaine de combinaisons de k parmi n éléments).

Au lieu des différentiations entières ($n \in \mathbb{N}$), on peut penser à une différentiation non entière ou "fractionnaire" $0 < \alpha < 1$ ou $n-1 < \alpha n$ et donc effectuer la différentiation d'ordre α, soit $\nabla^\alpha y_t = (1-L)^\alpha y_t$.

Cela ne pose pas de problème mathématique puisque le développement du binôme de Newton a déjà bien été généralisé à des puissances fractionnaires. En effet :

$\binom{n}{k} = \frac{n(n-1)...(n-k+1)}{k!}$. Donc $\binom{\alpha}{k} = \frac{\alpha(\alpha-1)...(\alpha-k+1)}{k!}$. Cela nous donne le coefficient de $(-L)^k$ dans le développement du binôme $(1-L)^\alpha$. Pour n entier, $\binom{n}{k}$ s'annule lorsque $k \geq n$. Mais, lorsque la puissance α est non entière, le kème coeffcient ne sera jamais nul et on obtient donc une infinité de coefficients :

$$(1-L)^\alpha y_t = \sum_{k=0}^{\infty} \binom{\alpha}{k} (-L)^k y_t$$

Notons que le numérateur du $k^{\text{ème}}$ coefficient $\alpha(\alpha-1)...(\alpha-k+1)$ est appelé fractionnelle décroissante ou descendante d'ordre k de α, notée $(\alpha)^k$. La fonction dfac(α, k, method="gamma") du package "betafunctions" de R nous permet de calculer les factorielles descendantes :

Ainsi $dfac(3, 2, ''gamma'')$ donne 6, le numérateur du $2^{\text{ème}}$ coefficient de $(1+x)^3$. Il suffit de le diviser par $2! = 2$ pour trouver ce coefficient $6/2 = 3$. En effet $(1+x)^3 = x^0 + 3x + 3x^2 + x^3$.

Pour $0 < \alpha < 1$, aucun facteur $(\alpha - k)$ dans la factorielle descendante $(\alpha)^k$ n'est nul et donc on aura une infinité de termes dans le développement du binôme de Newton $(1-L)^\alpha$.

FRACTIONAL

Pour $\alpha = 0.5$, les coefficients des premiers termes du développement seront :

- $k = 0$, $\binom{\alpha}{0} = \frac{(0.5)^0}{0!} = 1$

- $k = 1$, $\binom{\alpha}{1} = \frac{(0.5)^1}{1} = 0.5$

- $k = 2$, $\binom{\alpha}{2} = \frac{(0.5)^2}{2} = 0.25/2$

- $k = 3$, $\binom{\alpha}{3} = \frac{(0.5)^3}{3!} = 0.375/6$

Ainsi la demi différentiation d'une série chronologique y_t reviendra à multiplier cette série par $(1-L)^{0.5}$ et :
$\nabla^{0.5} y_t = (1-L)^{0.5} y_t = (L^0 - 0.5L - 0.125L^2 - 0.0625L^3 ...) y_t$.
Les 10 premiers coefficients de la différentiation d'ordre 0.5 sont ainsi :

```
 [1,]  1.00000000
 [2,] -0.50000000
 [3,] -0.12500000
 [4,] -0.06250000
 [5,] -0.03906250
 [6,] -0.02734375
 [7,] -0.02050781
 [8,] -0.01611328
 [9,] -0.01309204
[10,] -0.01091003
```

La différentiation d'ordre 1 revient à retrancher y_{t-1} de y_t. La différentiation d'ordre 0.5 revient à retrancher de y_t $0.5 y_{t-1}$, puis $0.125 y_{t-2}$, puis $0.0625 y_{t-3}$ et ainsi de suite jusqu'à l'infini. La différentiation d'ordre fractionnaire retranche donc de y_t des proportions décroissantes de toutes les valeurs passées de y_t.

3. Application de la différentiation fractionnaire aux séries temporelles

Nous venons de voir que la différentiation d'ordre fractionnaire de y_t consiste à retrancher de cette valeur une combinaison linéaire de toutes les valeurs passées y_{t-i}, i variant de 1 à l'infini. Les séries chronologiques économiques sont de taille finie. On ne peut donc pas leur appliquer en toute rigueur la différentiation fractionnaire qui implique de retrancher de chaque valeur des proportions d'une infinité de valeurs passées, lesquelles n'existent tout simplement pas. Même pas pour les valeurs les plus récentes. Quant aux dernières valeurs, le problème se pose même pour la différentation entière. On sait que la série des différences premières perd y_1, que la double différentiation perd y_1 et y_2.
La différntiation fractionnaire des séries chronologiques économiques ne peut donc être en pratique qu'une approximation. Cela tombe bien puisque les coefficients des y_{t-i} sont décroissants lorsque i augmente et très vite ils s'approchent de 0. On peut donc leur fixer un plancher (floor), en-deça duquel on néglige les termes correspondants dans la différentiation. C'est la solution qui est mise en œuvre par les différentes bibliothèques permettant d'effectuer la différentiation

fractionnaire, comme ”fracdiff” avec sa fonction ”diffseries(x, d)”.
En effet, lorsqu'on considère la série des coeffcients liés à la différentiation d'ordre 0.5 ci-dessus, on peut considérer qu'au-delà du 4ème terme, les coefficients sont négligeables. Ainsi, la différentiation d'ordre 0.5 revient à calculer la valeur suivante pour y_t :

$$\nabla^{0.5} y_t = (1-L)^{0.5} = y_t - 0.5y_{t-1} - 0.125y_{t-2} - 0.0625y_{t-3}$$

Mais quelle est la signification économique de la série obtenue en différenciant une série économique à l'ordre α, $0 < \alpha < 1$? On sait que la différentiation d'ordre 1 consiste à remplacer une série économique par ses variations absolues annuelles et que la différentiation d'ordre 2 consiste à remplacer une série par les variations des variations, ce qui peut être considérée comme l'accélération ou la décélération.
Lorsqu'on différencie d'ordre 0.5 par exemple, on retranche de y_t ses valeurs passées pondérées de coefficients diminuant au fur et à mesure que l'on remonte le temps. La signification économique est problématique. On peut cependant considérer que la différentiation fractionnaire consiste à remplacer une série par les différences entre les valeurs annuelles et les moyennes pondérées de toutes les valeurs antérieures. Il s'agit donc de substituer aux valeurs absolues les variations absolues par rapport à toutes les valeurs antérieures, au lieu de considérer les variations par rapport aux seules valeurs de $t-1$. En différenciant fractionnemment une série, on considère la variation d'une année par rapport à toutes les années antérieures. Il s'agit d'une variation ayant une plus grande signification que la variation par rapport à la seule année passée. En tous cas, avec la différentiation fractionnaire, on obtient une série stationnaire et conservant les caractéristiques de la série initiale. Ce qui est un double avantage.

4. Détermination de l'ordre fractionnaire d'intégration de séries

Maintenant que l'on sait comment effectuer la différentiation d'ordre α fractionnaire, on peut déterminer l'ordre fractionnaire ou entier d'intégration de n'importe quel série temporelle. Il suffit de commencer à tester l'intégration d'ordre 0 de la série considérée (en niveau) en effectuant un test ADF, par exemple sur la série. Si elle est stationnaire, alors elle est intégrée d'ordre 0, sinon, on peut la différentier d'ordre $\alpha = 0.001$ par exemple, ensuite tester si elle est stationnaire. Si elle l'est, elle est alors intégrée d'ordre $\alpha = 0.001$, sinon on augmente α de 0.001, on différentie de cet ordre et on teste et ainsi de suite jusqu'à ce que le test ADF donne comme résultat que la série est stationnaire, c'est-à-dire jusqu'à ce que sa p.value tombe en-dessous du seuil de Ficher, soit 0.05, nous permettant ainsi de rejeter l'hypothèse nulle de racine unitaire du test ADF. On aura ainsi trouvé l'ordre de différentiation permettant de rendre la série stationnaire, soit son ordre d'intégration.
En appliquant cet algorithme dans R aux séries du PIB, PIB agricole et PIB industriel du Maroc, on trouve que la série du PIB (en DH constants) pour la période 1966-2021 est intégrée d'ordre 1.265, la série du PIB agricole pour la même période est intégrée d'ordre 0.396, la série du PIB industriel pour la période 1967-2021 est intégrée d'ordre 0.275.
Ainsi pour le PIB, si on ignore la différentiation fractionnaire, il faudra différentier 2 fois la série pour la rendre stationnaire. Ce faisant, on perd l'information concernant les 2 premières années de la série et on perd aussi toute la mémoire de la série. La différentiation fractionnaire, permet de différentier une seule fois la série et d'effectuer ensuite sur les différences premières une différentiation d'ordre 0.265, ce qui permettra pour chaque année de tenir compte des valeurs du

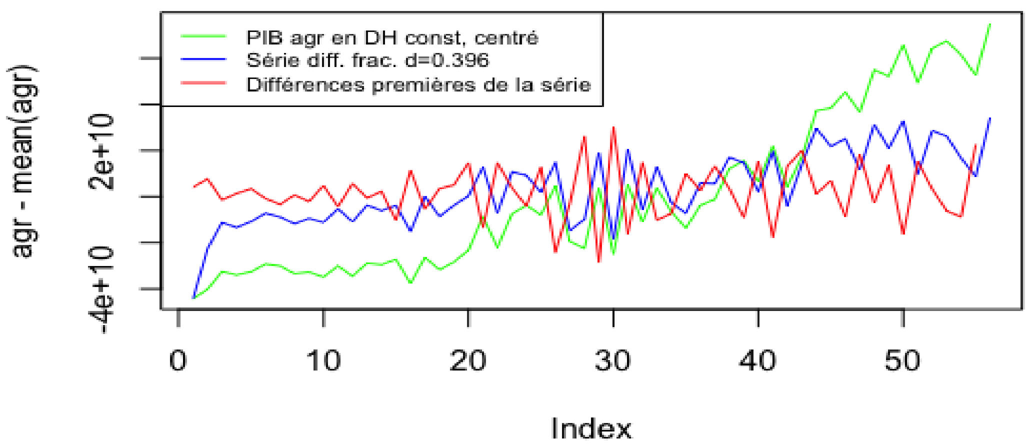
FRACTIONAL

PIB de toutes années antérieures (dans des proportions décroissantes bien évidemment).
Le gain de la différentiation et de l'intégration fractionnaires est encore plus clair et net avec la série du PIB agricole. Au lieu de considérer les premières différences de cette série comme série stationnaire, il suffit d'une différentiation d'ordre 0.396 de la série pour la rendre stationnaire. Cette différentiation d'ordre 0.396 consiste à retrancher de y_t la combinaison linéaire des valeurs passées de y_{t-i} pondérées par les coefficients suivants pour $i = 1, \ldots, 9$:

```
 [2,] -0.39600000
 [3,] -0.11959200
 [4,] -0.06394186
 [5,] -0.04162615
 [6,] -0.03000413
 [7,] -0.02302317
 [8,] -0.01843169
 [9,] -0.01521536
[10,] -0.01285529
```

5. Avantages de la différentiation fractionnaire

Nous venons de voir que la différentiation fractionnaire ou non entière tient compte de toutes les valeurs passées de la série. Elle conserve ainsi la mémoire de la série, alors que la différentiation entière efface cette mémoire. C'est là un avantage important de considérer l'intégration fractionnaire plutôt que l'intégration entière.
Représentons la série du PIB agricole en niveau, les premières différences de cette série et la série obtenue en différentiant d'ordre 0.396. On obtient :



Pour pouvoir représenter les 3 séries dans un même graphique, nous avons centré la série du PIB agricole.
Le graphique du PIB agricole en niveau est vert, celui des premières différences est rouge et celui de la série différentiée d'ordre 0.396 est bleu. On constate que les premières différences se sont écartées, plus que la série fractionnellement différentiée, de la série brute ou en niveau.
En effet, le coefficient de corrélation entre la série en niveau et la série différentiée à 0.396 est 0.8689, alors que ce coeffcient n'est que de 0.26 entre la série en niveau et les premières différences :

```
> cor(agr, agrfd)
[1] 0.8689069
> cor(agr[2:56],agr1d)
[1] 0.2624491
```

La différentiation fractionnaire permet donc de conserver les caractéristiques de la série brute, alors qu'avec la différentiation entière, les premières différences ne sont plus corrélées avec la série brute.
Cela améliore la qualité de prédiction de valeurs futures effectuées avec les modèles univariés et améliore la qualité des modèles (multivariés) de régression des séries chronologiques.
Pour la prédiction des valeurs futures du PIB agricole, on peut estimer un ARFIMA (autoregressive fractional integrated moving average model) de la série.
Notons tout de même que lorsque la série en différences premières est non stationnaire, c'est le cas du PIB marocain, alors il faut effectuer une différentiation d'ordre 1.265 (l'ordre d'intégration du PIB) et dans ce cas il faudra retrancher de y_t la valeur entière de y_{t-1} et des fractions des valeurs entières passées. La série qu'on obtient perd beaucoup des caractéristiques de la série en niveau et on trouve que la série du PIB en deuxièmes différences est plus corrélée (que la série différentiée à l'ordre 1.265) avec la série brute. La différentiation fractionnaire est donc plus utile lorsque l'ordre d'intégration des séries est compris entre 0 et 1.
Mais, pour l'analyse multivariée, régressons le PIB doublement différentié sur le PIB agricole différentié une fois. On obtient :

FRACTIONAL

TABLE 1

	Dependent variable :
	gdp2d
agr1d[2 :55]	0.374
	(0.247)
Constant	31,731,191,022.000***
	(3,044,531,237.000)
Observations	54
R^2	0.042
Adjusted R^2	0.024
Residual Std. Error	22,034,497,621.000 (df = 52)
F Statistic	2.293 (df = 1 ; 52)
Note :	*p<0.1 ; **p<0.05 ; ***p<0.01

En revanche, si on régresse le PIB fractionnellement différentié avec $d = 1.265$ sur le PIB agricole fractionnellement différencié avec $d = 0.396$, on obtient :

TABLE 2

	Dependent variable :
	gdpfd
agrfd	1.491***
	(0.413)
Constant	−269,504,815.000
	(6,804,640,531.000)
Observations	56
R^2	0.194
Adjusted R^2	0.179
Residual Std. Error	49,714,863,987.000 (df = 54)
F Statistic	13.005*** (df = 1 ; 54)
Note :	*p<0.1 ; **p<0.05 ; ***p<0.01

On voit bien que les deux modèles sont incomparables. Pour le premier qui est de mauvaise qualité, le PIB agricole n'a aucun effet significatif sur le PIB total. Pour le second modèle qui est assez bon avec R^2 ajusté de 0.18, le coefficient de l'agriculture est significatif et est de 1.491. Evidemment, on obtient de meilleurs résultats avec les taux de croissance qui sont stationnaires, donc intégrés d'ordre 0. Le modèle de régression du taux de croissance du PIB sur le taux de croissance du PIB agricole donne :

TABLE 3

	Dependent variable :
	cgdp
cagr	0.148^{***}
	(0.015)
Constant	0.037^{***}
	(0.003)
Observations	55
R^2	0.641
Adjusted R^2	0.634
Residual Std. Error	0.024 (df = 53)
F Statistic	94.583^{***} (df = 1 ; 53)
Note :	*p<0.1 ; **p<0.05 ; ***p<0.01

On obtient donc un modèle de bien meilleure qualité avec R^2 ajusté de 0.634 et le coeffcient de l'agriculture significatif de 0.148 et la constante (les variables non prises en compte dans le modèle) significative de 0.037. La croissance agricole explique donc près de 15% de la croissance économique du Maroc et les autres variables en expliquent 3.7%.

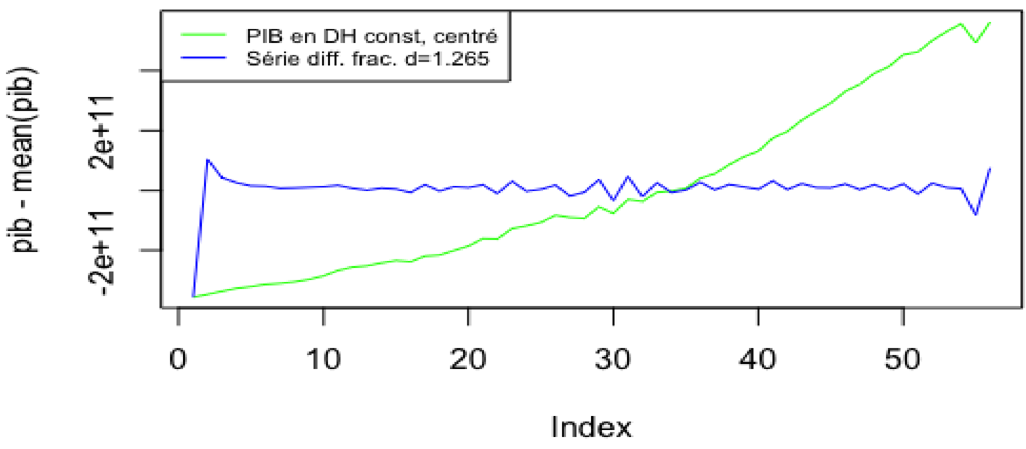
Les économètres considèrent que lorsque l'ordre d'intégration d'une série est $d < 0.5$, cette série est considérée comme intégrée d'ordre 0. Nous pouvons donc considérer que les séries des PIB agricole et industriel sont intégrées en niveau, puisque leurs ordres d'intégration sont respectivement de 0.396 et 0.275. En revanche, le PIB total est intégré d'ordre 1.265. Ses premières différences sont intégrées d'ordre 0.954 (la différence s'explique par la perte d'information causée par la différentiation première.
En tous cas en différenciant le PIB à l'ordre 1.265, on obtient une série stationnaire pouvant être régressée sur le PIB agricole. On obtient ainsi les mauvais résultats suivants :

FRACTIONAL

TABLE 4

	Dependent variable :
	pibfd
agr	0.253
	(0.215)
Constant	−11,715,451,025.000
	(16,001,185,252.000)
Observations	56
R^2	0.025
Adjusted R^2	0.007
Residual Std. Error	54,681,706,709.000 (df = 54)
F Statistic	1.385 (df = 1 ; 54)
Note :	*p<0.1 ; **p<0.05 ; ***p<0.01

La mauvaise qualité de ce modèle s'explique par le fait que la différentiation fractionnaire d'ordre $1 < d < 2$ donne une série stationnaire mais peu corrélée avec la série initiale :



En effet, le coefficient de corrélation entre les 2 séries n'est que de 0.107 ! Cela confirme notre hypothèse que la différentiation fractionnaire n'est intéressante que lorsque l'ordre d'intégration est inférieure à 1.