

CHAPITRE 4 :L'ANALYSE SPECTRALE **LES PROCESSUS ALÉATOIRES DANS LE DOMAINE** **DES FRÉQUENCES**

SECTION 1 : DÉFINITION

L'analyse spectrale a pour but d'examiner comment on peut exprimer une série temporelle comme superposition de fonctions cycliques élémentaires (i.e. $\sin \omega t$ et $\cos \omega t$). On peut alors examiner quelles sont les fréquences ω qui contribuent le plus à la dynamique de la série.

Le domaine de fréquence le plus choisis : $[0, \pi]$

SECTION 2 : GÉNÉRALITÉS **1. NOTIONS DE FRÉQUENCE ET DE PHASE**

Soit une fonction périodique de type :

$$\mathbf{SIN(\omega t + \phi)} \quad \text{ou} \quad \mathbf{COS(\omega t + \phi)}$$

ω est appelé **fréquence**

ϕ est appelé **phase**

La fréquence ω indique à quelle vitesse les cycles se déroulent. Augmenter la fréquence crée des cycles plus rapides : plus ω augmente est plus les cycles sont rapprochés.

Augmenter la phase ne fait que décaler les cycles.

2. DÉCOMPOSITION DE FOURIER

Opérer une décomposition de Fourier consiste à projeter une série sur une 'base', composée de fonctions élémentaires et orthogonales entre elles i.e. à trouver les coefficients a_j et b_j tels que :

$$y_t = \sum_{j=0}^{\infty} (a_j \cos \omega_j t + b_j \sin \omega_j t)$$

En pratique on ne mesure pas le processus y_t sur une période infinie mais sur T périodes, nous avons alors :

$$y_t = \sum_{j=0}^{L(t)} \left(a_j \cos \frac{2\pi j}{T} t + b_j \sin \frac{2\pi j}{T} t \right)$$

Avec $L(T) = T/2$ si T est pair ; $= (T-1)/2$ si T est impair.

Les coefficients a_j et b_j sont appelés coefficients de Fourier et le but est de les estimer par les OLS de la manière suivante :

$$\begin{aligned}\hat{a}_0 &= \frac{1}{T} \sum_{t=1}^T y_t = \bar{Y} \\ \hat{a}_j &= \frac{2}{T} \sum_{t=1}^T y_t \cos\left(\frac{2\pi j}{T}(t-1)\right) \\ \hat{b}_j &= \frac{2}{T} \sum_{t=1}^T y_t \sin\left(\frac{2\pi j}{T}(t-1)\right)\end{aligned}$$

Avec $j = 1, \dots, (T-1)/2$

3. LE PÉRIODOGRAMME

Le périodogramme se calcule de la manière suivante :

$$P(\omega_j) = \frac{T}{2} (\hat{a}_j^2 + \hat{b}_j^2)$$

que nous pouvons représenter avec les fréquences (ω) en abscisse (Cf. le corrélogramme).

La découverte de pics dans cette fonction indique que les fréquences pour lesquelles ces pics apparaissent sont les fréquences qui expliquent le mieux la dynamique de la série temporelle :

Les basses fréquences ($\omega \rightarrow 0$) indiquent des mouvements de long terme.

Les hautes fréquences ($\omega \rightarrow 3.14$) indiquent des mouvements de court terme.

Le périodogramme sert à repérer des périodicités dans les données

4. LE SPECTRE DE PUISSANCE OU DENSITÉ SPECTRALE

Soit l'autocovariance d'ordre j :

$$\sigma_j = E[(y_t - \mu)(y_{t-j} - \mu)]$$

En utilisant la transformée de Fourier on obtient la fonction de densité spectrale :

$$ds(\omega) = \frac{1}{2\pi} \left[\gamma_0 + 2 \sum_{j=1}^{\infty} \sigma_j \cos(\omega_j) \right]$$

5. DES MÉTHODES D'ESTIMATION DU SPECTRE

Il existe différentes méthodes d'estimation :

- Une estimation naturelle du spectre : en estimant la fonction d'autocovariance.**
- Une estimation par kernel, une fenêtre spectrale i.e. une fonction de pondération symétrique autour de 0.**
- Une estimation par lissage de la fonction d'autocovariance en se donnant une fonction de pondération appelée une fenêtre temporelle. Voici une liste des fenêtres les plus utilisées en pratique :**
 - Fenêtre rectangulaire,**
 - Fenêtre de Bartlett modifié,**
 - Fenêtre hamming,**
 - Fenêtre hanning,**
 - Fenêtre de Parzen...**

SECTION 3 : APPLICATIONS

- 1. Taux de croissance annuel du PIB belge**
- 2. La pression du sang dans les veines**
- 3. Les tâches solaires**
- 4. Le niveau du Nil**
- 5. Le taux de chômage...**

1. TAUX DE CROISSANCE ANNUEL DU PIB BELGE

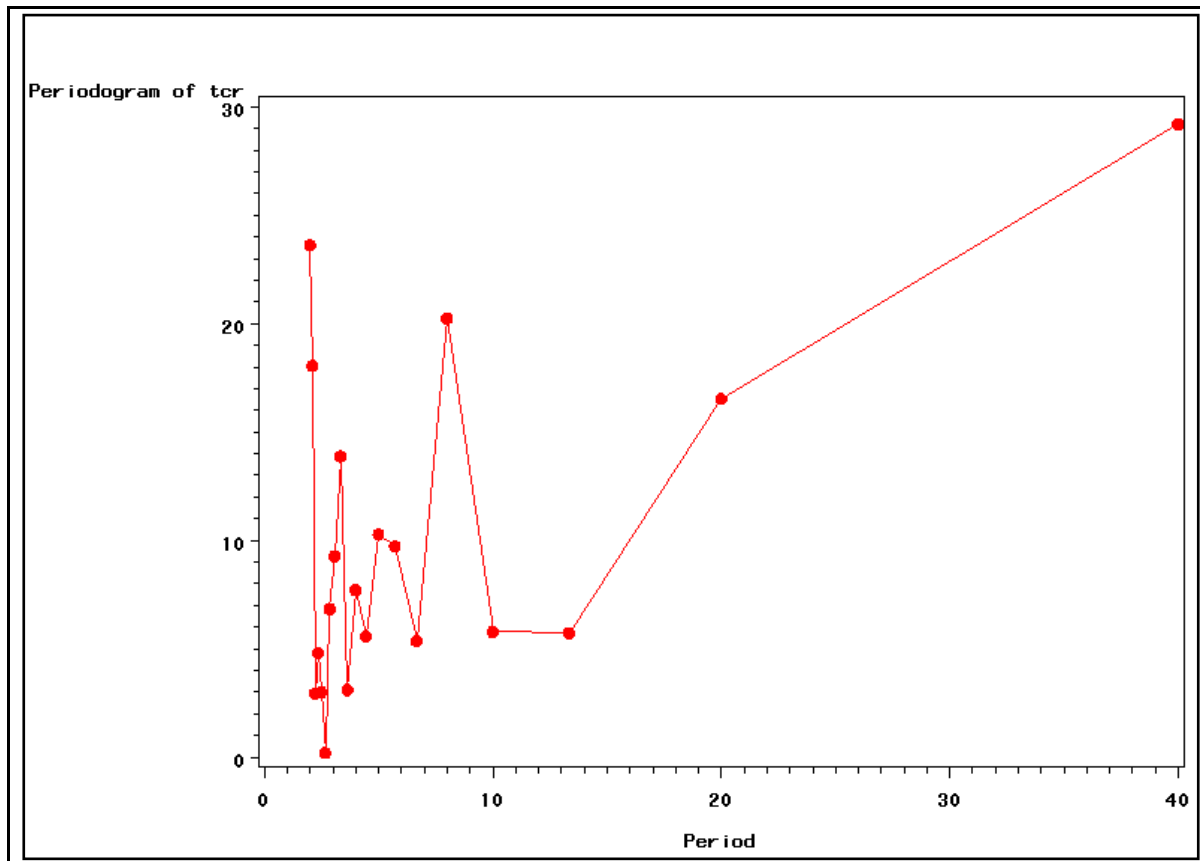
Les données du taux de croissance annuel du PIB belge proviennent du site internet Penn World Tables, mark 5.6 de 1950 à 1990 soit 40 données. Avec le logiciel SAS nous calculons les éléments suivants :

**Les fréquences,
Les périodes,
Les cosinus,
Les sinus,
Le périodogramme,
La densité spectrale...**

En utilisant différentes fenêtres...

Ensuite nous pouvons repérer par la série du périodogramme les périodes qui expliquent le mieux la dynamique :

Périodogramme (densité spectrale) du taux de croissance



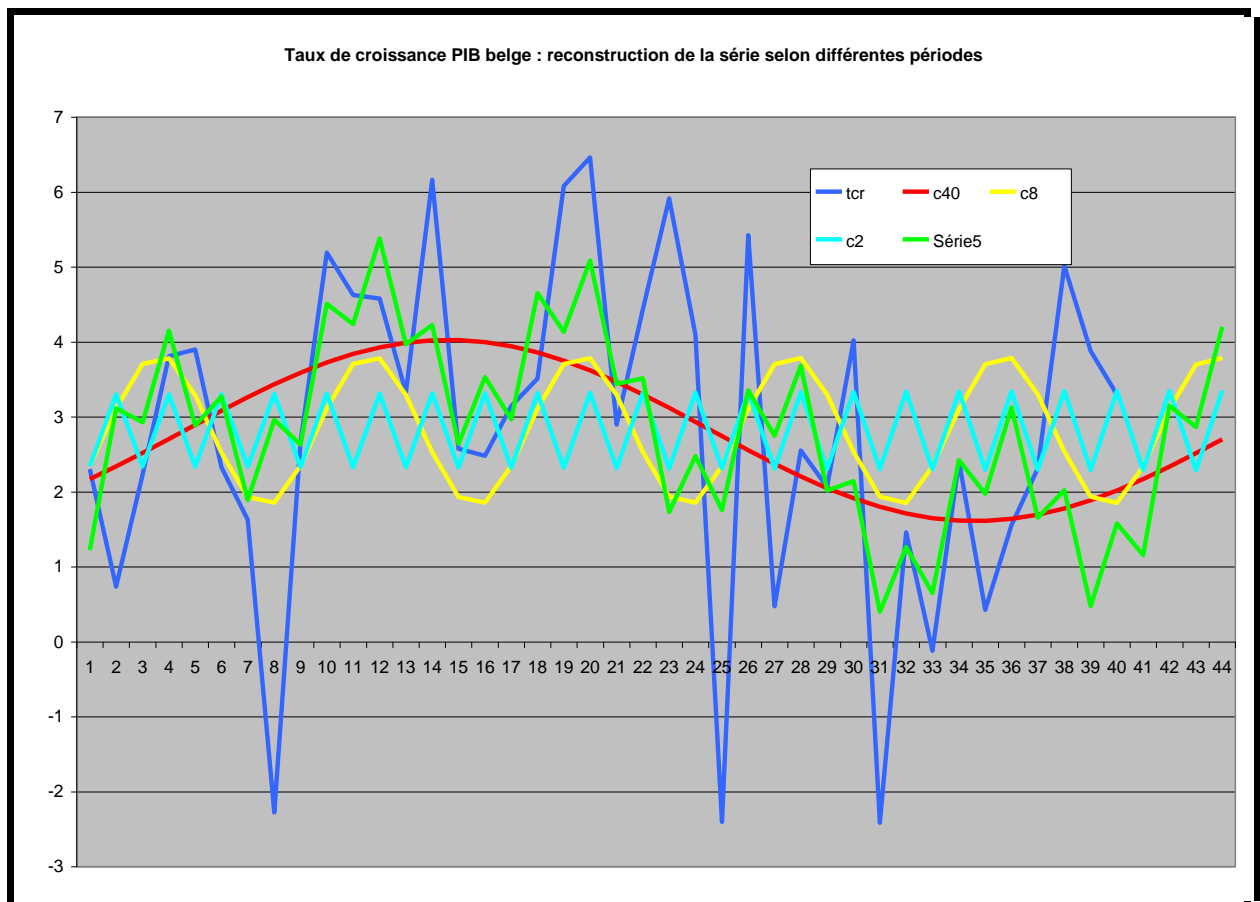
Mêmes éléments sous forme tableau :

PERIODES en année	PERIODOGRAMME
40	29,2017
20	16,5313
13,3333	5,7327
10	5,7925
8	20,2519
6,6667	5,3635
5,7143	9,7406
5	10,284
4,4444	5,5751
4	7,7165
3,6364	3,1128
3,3333	13,8834
3,0769	9,2812
2,8571	6,8535
2,6667	0,2015
2,5	3,0098
2,3529	4,8142
2,2222	2,9493
2,1053	18,0573
2	23,6265

On peut aussi le mettre en ordre décroissant et en % du total :

PERIODES	P_01	EN %	% CUMULE
40	29,2017	14,4577687	14,4577687
2	23,6265	11,6974858	26,1552545
8	20,2519	10,0267206	36,1819751
2,1053	18,0573	8,94017357	45,1221487
20	16,5313	8,18465061	53,3067993
3,3333	13,8834	6,87367468	60,1804739
5	10,284	5,09161087	65,2720848
5,7143	9,7406	4,8225734	70,0946582
3,0769	9,2812	4,59512435	74,6897826
4	7,7165	3,82044101	78,5102236
2,8571	6,8535	3,3931695	81,9033931
10	5,7925	2,86786814	84,7712612
13,3333	5,7327	2,83826115	87,6095224
4,4444	5,5751	2,76023335	90,3697557
6,6667	5,3635	2,65547014	93,0252259
2,3529	4,8142	2,38351158	95,4087374
3,6364	3,1128	1,54114803	96,9498855
2,5	3,0098	1,4901527	98,4400382
2,2222	2,9493	1,46019914	99,9002373
2,6667	0,2015	0,0997627	100

SAS nous fournit aussi les coefficients des cosinus et sinus de la décomposition de Fourier qui nous permettrons de reconstruire la série en tenant compte des différents cycles. Par exemple on pourrait reconstruire la série du PIB en utilisant uniquement la période de 40 ans, ou uniquement la période 8 ans mais aussi plusieurs compositions de périodes. Le graphique suivant nous en donne un aperçu :



La série 5 est une composition de période 40, 8 et 2 ans.

Donnons la démarche de reconstruction de la série C40 en utilisant les résultats de SAS et la formule :

$$y_t = \sum_{j=0}^{L(t)} \left(a_j \cos \frac{2\pi j}{T} t + b_j \sin \frac{2\pi j}{T} t \right)$$

Résultats de SAS

Obs	FREQ	PERIOD	COS_01	SIN_01	P_01	S_01
1	0.00000	.	0.00000	0.00000	0.0000	1.83828
2	0.15708	40.0000	-0.79834	0.90705	29.2017	1.72185
3	0.31416	20.0000	0.79761	-0.43634	16.5313	1.44418
4	0.47124	13.3333	0.18911	-0.50087	5.7327	1.11097
5	0.62832	10.0000	0.52640	-0.11193	5.7925	0.91442
6	0.78540	8.0000	-0.96142	0.29709	20.2519	0.85664
7	0.94248	6.6667	-0.45290	-0.25111	5.3635	0.77037
8	1.09956	5.7143	-0.21175	-0.66498	9.7406	0.75135
9	1.25664	5.0000	-0.46346	-0.54718	10.2840	0.67943
10	1.41372	4.4444	0.51940	0.09474	5.5751	0.60308
11	1.57080	4.0000	-0.26681	0.56093	7.7165	0.61815
12	1.72788	3.6364	0.39360	-0.02678	3.1128	0.61723
13	1.88496	3.3333	-0.53080	-0.64220	13.8834	0.63479
14	2.04204	3.0769	0.46475	0.49806	9.2812	0.58037
15	2.19911	2.8571	0.28780	-0.50975	6.8535	0.48530
16	2.35619	2.6667	0.07056	0.07137	0.2015	0.37510
17	2.51327	2.5000	0.04454	0.38537	3.0098	0.36820
18	2.67035	2.3529	-0.43858	-0.21990	4.8142	0.51791
19	2.82743	2.2222	0.38160	0.04296	2.9493	0.75570
20	2.98451	2.1053	0.47599	-0.82237	18.0573	1.01291
21	3.14159	2.0000	-1.08689	0.00000	23.6265	1.11545

Obs	tcr	c40	calculs
1	2,30330673	2,17274449	-0,79834*COS((2*3,14*1)/40)+0,90705*SIN((2*3,14*1)/40)+2,819443
2	0,73562193	2,34029373	
3	2,23500775	2,51962928	
4	3,80952381	2,70633978	
5	3,89908257	2,89583243	
6	2,32791491	3,08344602	
7	1,62777015	3,26456556	
8	-2,27711308	3,4347358	
9	2,66587678	3,5897708	
10	5,19330641	3,72585697	
11	4,62607424	3,8396468	
12	4,57881859	3,92834123	
13	3,32553476	3,98975853	
14	6,16205725	4,02238793	
15	2,5746496	4,02542679	

16	2,48032081	3,99880037	
17	3,14492754	3,94316364	
18	3,51271603	3,85988516	
19	6,0811728	3,75101346	
20	6,46193218	3,6192266	
21	2,89663462	3,46776634	
22	4,43873379	3,30035836	
23	5,91656414	3,12112062	
24	4,09714889	2,9344621	
25	-2,40413877	2,74497428	
26	5,42563143	2,55731828	
27	0,47323277	2,37611013	
28	2,55127073	2,20580727	
29	2,06678787	2,05059886	
30	4,02174932	1,91430281	
31	-2,41528479	1,80027176	
32	1,4591799	1,7113107	
33	-0,11833242	1,64960794	
34	2,4059054	1,61668126	
35	0,42716027	1,6133406	
36	1,55959238	1,63966815	
37	2,31218916	1,69501629	
38	5,0144977	1,77802353	
39	3,8817606	1,88664804	
40	3,29893684	2,01821782	
41		2,16949647	
42		2,33676277	
43		2,51590223	
44		2,70250832	$=-0,79834 \cdot \cos((2 \cdot 3,14 \cdot 44)/40) + 0,90705 \cdot \sin((2 \cdot 3,14 \cdot 44)/40) + 2,819443$

References

- Bard, J. (1970), "Comparison of Gradient Methods for the Solution of the Nonlinear Parameter Estimation Problem," *SIAM Journal of Numerical Analysis*, 7, 157–186.
- Bard, J. (1974), *Nonlinear Parameter Estimation*, New York: Academic Press.
- Bates, D. M., and Watts, D. L. (1980), "Relative Curvature Measures of Nonlinearity (with Discussion)," *Journal of the Royal Statistical Society, Series B*, 1–25.
- Bates, D. M., and Watts, D. L. (1981), "A Relative Offset Orthogonality Convergence Criterion for Nonlinear Least Squares," *Technometrics*, 123, 179–183.
- Box, M. J. (1971), "Bias in Nonlinear Estimation (with Discussion)," *Journal of the Royal Statistical Society, Series B*, 171–201.
- Beaton, A. E. and Tukey, J. W. (1974), "The Fitting of Power Series, Meaning Polynomials, Illustrated on Band-Spectroscopic Data," *Technometrics*, 16, 147–185.
- Charnes, A., Frome, E. L., and Yu, P. L. (1976), "The Equivalence of Generalized Least Squares and Maximum Likelihood Estimation in the Exponential Family," *Journal of the American Statistical Association*, 71, 169–172.
- Cook, R. D. and Tsai, C. L. (1985), "Residuals in Nonlinear Regression," *Biometrika*, 72, 23–9.
- Cox, D. R. (1970), *Analysis of Binary Data*, London: Chapman & Hall.
- Finney, D. J. (1971), *Probit Analysis*, Third Edition, Cambridge: Cambridge University Press.
- Gallant, A. R. (1975), "Nonlinear Regression," *American Statistician*, 29, 73–81.
- Gill, P. E., Murray, W., and Wright, M. H. (1981), *Practical Optimization*, New York: Academic Press.
- Goodnight, J. H. (1979), "A Tutorial on the Sweep Operator," *American Statistician*, 33, 149–158.
- Hartley, H. O. (1961), "The Modified Gauss-Newton Method for the Fitting of Non-Linear Regression Functions by Least Squares," *Technometrics*, 3, 269–280.
- Holland, P. H. and Welsch, R. E. (1977), "Robust Regression Using Iteratively Reweighted Least-Squares," *Communications Statistics: Theory and Methods*, 6, 813–827.
- Hougaard, P. (1982), "Parameterizations of Nonlinear Models," *Journal of the Royal Statistical Society, Series B*, 244–252.
- Hougaard, P. (1985), "The Appropriateness of the Asymptotic Distribution in a Nonlinear Regression Model in Relation to Curvature," *Journal of the Royal Statistical Society, Series B*, 103–114.
- Huber, P. J. (1964), "Robust Estimation of a Location Parameter," *Annals of Mathematical Statistics*, 35, 73–101.
- Huber, P. J. (1973), "Robust Regression: Asymptotics, Conjectures, and Monte Carlo," *Annals of Statistics*, 1, 799–821.
- Jennrich, R. I. (1969), "Asymptotic Properties of Nonlinear Least Squares Estimators," *Annals of Mathematical Statistics*, 40, 633–643.
- Jennrich, R. I. and Moore, R. H. (1975), "Maximum Likelihood Estimation by Means of Nonlinear Least Squares," *American Statistical Association, 1975 Proceedings of the Statistical Computing Section*, 57–65.
- Jennrich, R. I. and Sampson, P. F. (1968), "Application of Stepwise Regression to Nonlinear Estimation," *Technometrics*, 10, 63–72.

Judge, G. G., Griffiths, W. E., Hill, R. C., and Lee, T.-C. (1980), *The Theory and Practice of Econometrics*, New York: John Wiley & Sons.

Kennedy, W. J. and Gentle, J. E. (1980), *Statistical Computing*, New York: Marcel Dekker.

Lee, E. T. (1974), "A Computer Program for Linear Logistic Regression Analysis," *Computer Programs in Biomedicine*, 80–92.

Marquardt, D. W. (1963), "An Algorithm for Least-Squares Estimation of Nonlinear Parameters," *Journal for the Society of Industrial and Applied Mathematics*, 11, 431–441.

McCullagh, P. and Nelder, J. A. (1989), *Generalized Linear Models*, Second Edition, New York: Chapman & Hall.

Nelder, J. A. and Wedderburn, R. W. M. (1972), "Generalized Linear Models," *Journal of the Royal Statistical Society, Series A*, 135, 370–384.

Pinheiro, J. C. and Bates, D. M. (1995), "Approximations to the Log-Likelihood Function in the Nonlinear Mixed-Effects Model," *Journal of Computational and Graphical Statistics*, 4, 12–35.

Pringle, R. M. and Rayner, A. A. (1971), *Generalized Inverse Matrices with Applications to Statistics*, New York: Hafner Publishing Co.

Ratkowsky, D. (1983), *Nonlinear Regression Modeling*, New York and Basel: Marcel Dekker.

Ratkowsky, D. (1990), *Handbook of Nonlinear Regression Models*, New York and Basel: Marcel Dekker.

Schabenberger, O. and Pierce, F. J. (2002), *Contemporary Statistical Models for the Plant and Soil Sciences*, Boca Raton, FL: CRC Press.

Schabenberger, O., Tharp, B. E., Kells, J. J., and Penner, D. (1999), "Statistical Tests for Hormesis and Effective Dosages in Herbicide Dose Response," *Agronomy Journal*, 91, 713–721.

Seber, G. A. F. and Wild, C. J. (1989) *Nonlinear Regression*, New York: John Wiley & Sons.

St. Laurent, R. T. and Cook, R. D. (1992), "Leverages and Superleverages in Nonlinear Regression," *Journal of the American Statistical Association*, 87, 985–990.

St. Laurent, R. T. and Cook, R. D. (1993), "Leverages, Local Influence, and Curvature in Nonlinear Regression," *Biometrika*, 80, 99–106.