

K. EICHLER und V. WAGENITZ

Berechnung von "Fourier-Koeffizienten" bei nicht-äquidistanten Frequenzen

Entwickelt man eine digitalisiert vorliegende Zeitreihe der Länge T in eine Fourier-Reihe (z.B. mithilfe der FFT), erhält man Fourier-Koeffizienten bei diskreten Frequenzen $\omega_k = k \cdot \omega_0 = k \cdot \frac{2\pi}{T}$. Werden jedoch Funktions- oder Meßwerte nicht in Abhängigkeit von der Frequenz ω , sondern z.B. von $\log T$ (wie in der MT üblich) benötigt, entsteht hierbei eine starke Häufung der Fourier-Koeffizienten zu kurzen Perioden hin, die u.U. eine frequenzabhängige Mittelung z.B. der Powerspektren nötig macht.

Das im folgenden skizzierte Verfahren erlaubt es, Entwicklungskoeffizienten trigonometrischer Funktionen für beliebige, in Grenzen frei vorgegebene Frequenzen zu bestimmen, die formal als "Fourier-Koeffizienten" bei nicht-äquidistanten Frequenzen angesehen werden können, ohne daß diese Approximation jedoch einer vollständigen Entwicklung nach orthogonalen Funktionssystemen entspricht.

Die Zeitreihe $y(t)$ sei zu M nicht notwendig gleichabständigen Zeitpunkten t_j bekannt: $y_j := y(t_j)$, $j=1 \dots M$. Sie soll durch eine Summe trigonometrischer Funktionen mit N frei vorgegeben, nicht notwendig gleichabständigen Frequenzen ω_k , $k=1 \dots N$ im Sinne der kleinsten Fehlerquadratsumme angepaßt werden:

$$y_j + e_j = \frac{a_0}{2} + \sum_{k=1}^N (a_k \cos \omega_k t_j + b_k \sin \omega_k t_j) \quad (1)$$

$$\text{und } \sum_{j=1}^M e_j^2 =: \min. \quad j=1 \dots M$$

Zu bestimmen sind die $2N + 1$ Koeffizienten $a_0, a_1 - a_N, b_1 - b_N$. Für die Lösung dieser Ausgleichsaufgabe in der üblichen Art (z.B. nach Zurmühl, [1958]) müssen möglichst viele, mindestens aber $M > 2N + 1$ Datenpaare (t_j, y_j) zur Verfügung stehen.

Das lineare Gleichungssystem (1) lautet in Matrixschreibweise:

$$\underline{T} \underline{c} = \underline{y} + \underline{e} \quad \text{oder} \quad \underline{T} \underline{c} - \underline{y} = \underline{e} \quad (\text{Fehlergleichungen}).$$

Dabei ist

$$\underline{\underline{T}} = \begin{pmatrix} 1/2 & \cos \omega_1 t_1 & \cdots & \cos \omega_N t_1 & \sin \omega_1 t_1 & \cdots & \sin \omega_N t_1 \\ \cdot & \cdot & & \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot & \cdot & & \cdot \\ 1/2 & \cos \omega_1 t_M & \cdots & \cos \omega_N t_M & \sin \omega_1 t_M & \cdots & \sin \omega_N t_M \end{pmatrix}$$

eine Matrix mit M Zeilen und $2N + 1$ Spalten, $M > 2N + 1$. Rang $\underline{\underline{T}} = 2N + 1$, d.h. die Matrix ist (außer bei extrem pathologischer Wahl der ω_k und der t_j) spaltenregulär, so daß die Spalten der Matrix linear unabhängig sind.

$$\underline{c} = (a_0, a_1 \cdots a_N, b_1 \cdots b_N)^t$$

ist der Vektor der gesuchten $2N + 1$ Koeffizienten,

$$\underline{y} = (y_1 \cdots y_M)^t \quad \text{und} \quad \underline{e} = (e_1 \cdots e_M)^t$$

sind die Vektoren der Funktionswerte sowie der Anpassungsfehler.

Das Gleichungssystem $\underline{\underline{T}} \underline{c} = \underline{y}$ ist überbestimmt, da die Daten i.a. fehlerbehaftet sind ($\underline{e} \neq 0$; mehr Gleichungen als Unbekannte). Mithilfe der Gauß'schen Transformation wird es in das konsistente System der Normalgleichungen übergeführt:

$$\underline{\underline{T}}^t \underline{\underline{T}} \underline{c} = \underline{\underline{T}}^t \underline{y} \quad (2)$$

Die quadratische Matrix $\underline{\underline{T}}^t \underline{\underline{T}}$ (Dimension: $(2N+1) \times (2N+1)$) ist nicht-singulär (d.h. invertierbar), symmetrisch, positiv definit und von kleineren Ausmaßen als die Ausgangsmatrix $\underline{\underline{T}}$. Die gewünschten Entwicklungskoeffizienten der trigonometrischen Funktionen stehen nach Auflösung des Gleichungssystems (2) im Vektor \underline{c} zur Verfügung. Will man das gleiche Raster der Frequenzen ω_k und relativen Zeiten t_j für mehrere Datensätze anwenden, lohnt es sich, nicht das Gleichungssystem (2) zu lösen, sondern die Matrix $\underline{\underline{T}}^t \underline{\underline{T}}$ zu invertieren:

$$\underline{c} = (\underline{\underline{T}}^t \underline{\underline{T}})^{-1} \underline{\underline{T}}^t \underline{y} = \underline{\underline{S}} \underline{y} \quad \text{mit} \quad \underline{\underline{S}} = (\underline{\underline{T}}^t \underline{\underline{T}})^{-1} \underline{\underline{T}}^t \quad (3)$$

Die nur von den ω_k und t_j abhängige Matrix $\underline{\underline{S}}$ braucht dabei nur einmal berechnet zu werden und liefert dann nach Anwendung auf beliebig viele Datensätze die jeweils dazugehörigen Vektoren \underline{c} der Entwicklungskoeffizienten. Die Lösung des Gleichungssystems (3) ist formal ähnlich der Bestimmung der Übertragungsfunktionen eines linearen Systems mit $2N+1$ Eingängen, wenn man $\underline{\underline{T}}^t \underline{y}$ als Vektor der "Kreuzspektren" zwischen "Ein- und Ausgang" und $\underline{\underline{T}}^t \underline{\underline{T}}$ als Matrix der "Kreuz- und Quad-Spektren" der "Eingänge" untereinander auffaßt.

Als Beispiel wurde obiges Schema benutzt, um aus einer diskreten Zeitreihe $y_j = y(t_j)$, $\Delta t = 1$ sec, die einem AR Prozeß 2. Ordnung entspricht (nach Jenkins & Watts, [1968], pp. 164, 221, 228-9, mit $\alpha_1 = 1.0$, $\alpha_2 = -0.5$, $\mu = 0$, $\sigma = 1.0$), Schätzwerte des zugehörigen Autopowerspektrums P_{yy} für die folgenden acht fest vorgegebenen Perioden $T_k = \frac{2\pi}{\omega_k}$ zu berechnen:

k	0	1	2	3	4	5	6	7	8
T_k in sec	∞	64	40	25	15.5	10	6.4	4	2.5

Die Approximation mit trigonometrischen Funktionen nach Gleichung (1) wurde über jeweils $M = 64$ Werte durchgeführt. Die P_{yy} - Schätzwerte wurden durch Mittelung der quadrierten (a_k, b_k) - Koeffizienten über 16 solcher Datensätze gewonnen.

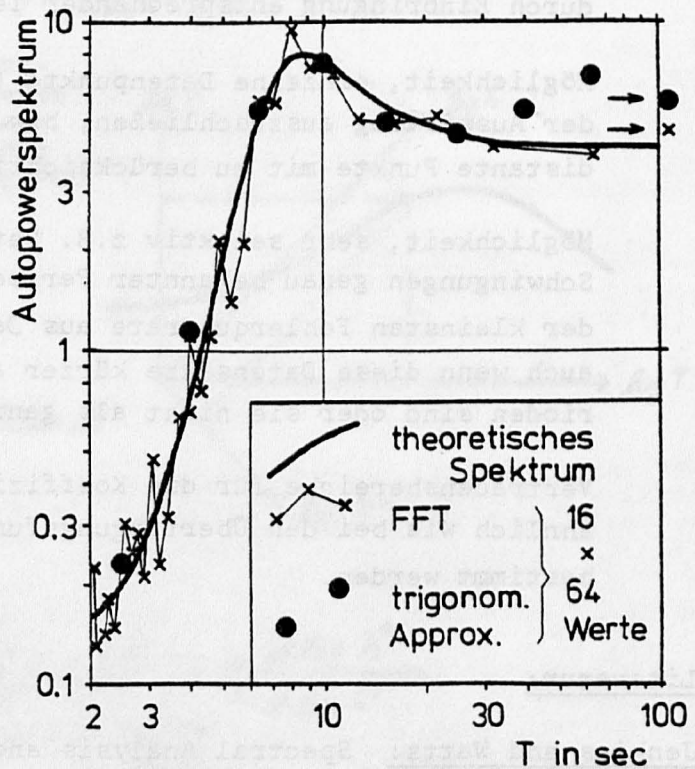
Abb. 1: Autopowerspektrum P_{yy} eines autoregressiven Prozesses 2. Ordnung:

$$y_j = \alpha_1 y_{j-1} + \alpha_2 y_{j-2} + z_j$$

mit $\alpha_1 = 1.0$, $\alpha_2 = -0.5$

und z_j normalverteilt

mit $\mu = 0$, $\sigma = 1.0$



Als Ergebnis zeigt Abbildung 1 die acht mit der so durchgeführten trigonometrischen Approximation berechneten $P_{yy}(T_k)$ - Schätzwerte zusammen mit dem theoretischen Spektrum und den aus dem gleichen Datenmaterial und gleicher Unterteilung durch Anwendung der FFT berechneten Schätzwerten. Zu erkennen ist eine gute Übereinstimmung

zwischen den $P_{yy}(T_k)$ - Schätzwerten und dem theoretischen Spektrum für $T < 40$ sec. Ob die größeren Abweichungen bei $T = 40$ und 64 sec zufällig oder systematisch (und, falls systematisch: wodurch bedingt) sind, wurde nicht untersucht. Die mithilfe der FFT gewonnenen Spektralschätzwerte zeigen die übliche Häufung bei kurzen Perioden; eine Glättung oder Mittelung im Frequenzbereich würde das theoretische Spektrum offensichtlich ebenfalls gut annähern.

Grundsätzlich scheint es demnach zumindest bei der Berechnung von Powerspektren möglich zu sein, mit dem angegebenen Verfahren vernünftige Schätzwerte bei fest vorgegebenen, nichtäquidistanten Frequenzen zu erhalten.

Mögliche Vorteile des Verfahrens könnten sein:

einfache Eliminierung von linearen, quadratischen o.ä. Trends durch Einbringung entsprechender Terme in Gleichung (1).

Möglichkeit, einzelne Datenpunkte (z.B. Ausreißer o.ä.) von der Auswertung auszuschließen, bzw. auch zeitlich nicht-äquidistante Punkte mit zu berücksichtigen.

Möglichkeit, sehr selektiv z.B. Beträge und Phasen mehrerer Schwingungen genau bekannter Periodendauer optimal im Sinne der kleinsten Fehlerquadrate aus Datensätzen zu bestimmen... auch wenn diese Datensätze kürzer als die entsprechenden Perioden sind oder sie nicht als ganze Vielfache beinhalten.

Vertrauensbereiche für die Koeffizienten (a_k, b_k) können ähnlich wie bei den Übertragungsfunktionen linearer Systeme bestimmt werden.

Literatur:

Jenkins and Watts: Spectral Analysis and its Applications. Holden-Day, San Francisco, 1968.

Zurmühl: Matrizen. Springer, Berlin, 1958.