

Supplemento Statistico - Matematico

al Vol. VII

**dei "Rendiconti del Comitato Veneto
per il potenziamento degli studi economici
e per la programmazione".**

Gennaio 1974 - Venezia

ANALISI SPETTRALE DELLE SERIE
DELL' INDICE DEL COSTO DELLA
VITA (1951 - 1971)

di

Domenico Sartore

Supplemento Statistico - Matematico al Vol. VII dei "Rendiconti del Comitato Veneto per il potenziamento degli studi economici e per la programmazione" .

Premessa

Lo spunto per questa nota è stato offerto dall'idea di analizzare le interrelazioni cicliche tra prezzi, salari e produzione nei diversi settori industriali italiani dal 1951 al 1971. L'indagine doveva essere condotta tramite l'Analisi Spettrale, cioè con una strumentazione matematico-statistica di recente applicazione a fenomeni di tipo economico.

Nell'iniziare tale ricerca si è vista l'opportunità di puntualizzare alcuni aspetti riguardanti l'Indice del Costo della Vita elaborato dall'ISTAT, sottolineandone la carenza nell'interpretazione dell'andamento ciclico dei prezzi al consumo per le famiglie del ceto operaio-impiegatizio dei settori non agricoli.

E' noto come già da diverso tempo siano state mosse delle critiche non in differenti ai criteri di computo di tale indice. In questa sede si è ricevuta conferma della rigidità dell'indice generale nel rappresentare la variabilità ciclica dovuta alla dinamica dei prezzi per i diversi Capitoli di spesa considerati dall'ISTAT.

Le esigenze del presente lavoro non permettono di dare nemmeno succintamente una rigorosa impostazione metodologica dell'analisi spettrale data la complessità e la vastità di tale teoria. Esiste, tuttavia una ricca bibliografia in lingua inglese che è riportata per i testi più significativi in appendice.

§.1 Osservazioni generali sulle serie storiche economiche trattate con l'Analisi Spettrale.

L'Analisi Spettrale si è sviluppata metodologicamente nello studio delle serie temporali stazionarie. Facilmente definibile analiticamente nel caso di processi stocastici di cui si conosca la funzione di distribuzione, la stazionarietà diventa semplice intuizione nel caso di fenomeni economici di cui si dispone inevitabilmente di singoli campioni unitari. Generalmente ci si riferisce alla particolare classe delle serie stazionarie di secondo ordine caratterizzata dalla costanza sull'intero asse temporale della loro media e varianza. L'irreale ipotesi di estensibilità all'infinito del numero di campioni estraibili e, almeno nel caso economico, l'unicità delle loro determinazioni, comporta l'evidente inapplicabilità operativa di tale caratterizzazione: per una serie temporale con un numero finito di dati è impossibile distinguere l'attribuzione dell'andamento evolutivo ad un trend di tipo polinomiale piuttosto che a ciclicità di frequenza molto bassa, relativa quindi ad un periodo temporale superiore al campo di rilevazione della serie.

In contrapposizione alle serie stazionarie, le serie evolutive non risultano omogenee nella loro struttura rispetto al tempo: la presenza di un trend in media è l'aspetto più evidente di non stazionarietà. È importante notare la relazione esistente tra trend e l'intervallo di tempo della serie temporale. Un "segmento" di variazione stagionale può essere concettualmente considerato trend all'interno di un periodo di tempo abbastanza breve.

Nello studio dei problemi economici, le serie temporali si presentano generalmente con andamenti ciclici attorno a trends in media che riflettono la graduale crescita estensiva e tecnologica del processo economico globale; esse vengono rilevate istantaneamente o in modo cumulato rispetto al tempo. In quest'ultimo caso dovranno considerarsi gli effetti dovuti alle variazioni di calendario, mentre per le serie rilevate istantaneamente

sorge il problema cosiddetto di "Aliasing", riferito alla possibilità di confondere, nell'analisi, l'importanza dei vari contributi frequenziali alla variabilità del fenomeno qualora l'intervallo di registrazione fissato (unità di misura temporale) escluda possibili ciclicità esistenti all'interno del medesimo.

La Statistica Economica tradizionale, quando analizza l'andamento di un fenomeno economico, propone il seguente modello interpretativo:

$$G = f(T,C,S,A)$$

dove

G è la serie dei dati grezzi

T è la componente "Trend"

C è la componente ciclica

S è la componente di stagionalità

A è la componente accidentale

L'applicazione dell'analisi spettrale ad una serie temporale stazionaria comporta anche essa una scomposizione del fenomeno non più nel dominio temporale ma frequenziale che presenta qualche analogia con il modello citato, dove T è identicamente 0 oppure 1 a seconda dell'additività o moltiplicabilità di tale modello.

Per una precisa comprensione della scomposizione frequenziale e dell'analisi spettrale è opportuno accennare al significato della stima spettrale di una serie temporale, la cui visualizzazione grafica costituisce lo spettrogramma, evitando l'intera costruzione teorica che porta alla definizione di tale stima.

§.2 Cenni sulla Teoria dell'Analisi Spettrale.

Data una serie temporale x_t definita come successione di valori in $t = 1, 2, \dots, n$ di essa si può dare una rappresentazione grafica sul dominio temporale. In generale si può supporre che $x(t)$ sia realmente una funzione continua del tempo, sebbene registrata per intervalli regolari. La rappresentazione così data del fenomeno può condurre a rilevare l'eventuale andamento evolutivo e al più un andamento oscillatorio attorno al trend. La descrizione fatta dall'analisi spettrale nel dominio frequenziale di un fenomeno stazionario o comunque depurato dalla componente evolutiva equivale, in prima approssimazione, ad interpretarlo come una scomposizione in componenti cicliche regolari con l'indicazione della loro importanza nella caratterizzazione del suo andamento. Da un punto di vista matematico è dimostrabile che, qualora siano soddisfatte particolari condizioni, una funzione deterministica periodica di periodo qualsiasi T , cioè una funzione per cui vale

$$x(t + T) = x(t)$$

è esprimibile in serie di Fourier, cioè

$$\begin{aligned} x(t) &= \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos 2\pi nft + b_n \sin 2\pi nft) = \\ &= \frac{a_0}{2} + \sum_{n=1}^{\infty} A_n \cos (2\pi nft + \theta_n) \end{aligned}$$

dove a_n e b_n sono detti coefficienti di Fourier della serie.

f è la frequenza fondamentale (espressa in cicli per unità di tempo) legata al periodo della serie mediante la semplice relazione $f = \frac{1}{T}$.

La frequenza fondamentale angolare sarà allora $2\pi f$

θ_n indica la fase dell'armonica ennesima, cioè il valore di t nel quale l'onda armonica raggiunge il suo vertice

A_n indica l'ampiezza cioè l'altezza massima o depressione minima dell'on-
da cosinusoidale misurata dal suo livello medio $-\frac{A_0}{2}$ (con $A_0 = a_0$)

ogni singolo termine della sommatoria è denominato armonica di ordine n
della serie. Giova notare come A_n e θ_n siano ottenibili dai coefficien-
ti di Fourier mediante la loro semplice tramutazione in coordinate polari e
quindi valgono le seguenti relazioni:

$$a_n = A_n \cos \theta_n$$

$$-b_n = A_n \sin \theta_n$$

per cui

$$A_n = \sqrt{a_n^2 + b_n^2} \quad \text{e} \quad \theta_n = \arctg \left(-\frac{b_n}{a_n} \right)$$

Lo sviluppo in serie di Fourier della $x(t)$ (supposta periodica) ammet-
te una rappresentazione più generale nel campo complesso, sintetizzata dal-
la seguente espressione

$$(1) \quad a(t) = \sum_{-\infty}^{\infty} F(n) e^{i n 2 \pi f t}$$

$$F(n) = \frac{1}{2} (a_n - i b_n) \quad F(n) \text{ essendo denominata}$$

funzione complessa
dell'ampiezza
("complex amplitude")

ed

$$F(0) = \frac{1}{2} a_0$$

Tale estensione permette di interpretare l'ampiezza complessa, attraverso
la relazione esistente con i coefficienti di Fourier, come trasformata di
Fourier della funzione periodica $x(t)$, cioè :

$$(2) \quad F(n) = \frac{1}{T} \int_0^T x(t) e^{-i n 2 \pi f t} dt$$

$n=0, \pm 1, \pm 2, \dots$

L'espressione (2) consiste in una funzione dell'ordine n delle armoniche ed è una rappresentazione della funzione periodico-temporale nel dominio di frequenza. Il significato di $F(n)$ traspare ora più evidente; si osservi che moltiplicando $x(t)$ con il fattore $e^{-in2\pi ft}$ ed assumendo il valor medio del prodotto sopra il periodo, per qualsiasi valore di n si ottiene un'analisi di $x(t)$ relativamente all'ampiezza ed alla fase delle cosinusoidi in cui la $x(t)$ stessa è esplicitata. Ciò è più evidente se si considera che :

$$F(n) = \frac{1}{2} (a_n - i b_n) = \frac{1}{2} (A_n \cos \theta_n + i A_n \sin \theta_n) = \frac{1}{2} A_n e^{i \theta_n}$$

Le informazioni concernenti le ampiezze e le fasi delle armoniche componenti la $x(t)$ devono ritenersi esaustive in quanto $F(x)$ è univocamente determinata dalla (1). L'espressione (1), a sua volta trasforma

ta inversa di Fourier della funzione $F(n)$, interpreta la funzione periodica $x(t)$ come somma di cosinusoidi di cui $F(n)$ rappresenta appunto sia l'ampiezza che la fase, in questo senso si può affermare che la (1) è una espressione *sintetica* in contrapposizione alla (2) che invece è un'espressione *analitica* (si veda [9], pag. 7). La *funzione spettro*, da alcuni denominata *spettro potenza*, è il prodotto della funzione complessa dell'ampiezza per la sua coniugata e viene indicata con $\Phi(n)$, cioè :

$$\Phi(n) = |F(n)|^2 = \frac{1}{4} A_n^2$$

Quindi nel caso di funzioni periodiche, $\Phi(n)$ misura l'importanza delle diverse armoniche componenti la serie fornendo informazioni circa la loro ampiezza mentre trascura completamente le informazioni relative alla loro fase.

L'importanza che ha assunto l'Analisi Spettrale è legata all'estensione del concetto di funzione spettrale sia alla classe delle funzioni aperiodiche che alla classe delle funzioni casuali. Nel caso di funzioni aperiodiche grazie ad un teorema dovuto a Plancherel (si veda [11], pag. 436) che assicura l'esistenza della funzione $F(f)$, integrabile secondo Riemann^(x),

(x) Più generalmente, e ciò potrebbe avere anche interesse, la funzione $F(f)$ deve essere pensata come elemento della classe L_2 delle funzioni integrabili secondo Lebesgue

è possibile definire analogamente alle funzioni periodiche:

$$(3) \quad x(t) = \int_{-\infty}^{\infty} F(f) e^{i2\pi ft} df$$

dove

$$F(f) = \int_{-\infty}^{\infty} x(t) e^{-i2\pi ft} dt$$

L'espressione (3) viene denominata *integrale di Fourier* per una funzione aperiodica ed indica che la funzione aperiodica $x(t)$ è sintetizzata da un aggregato infinito di sinusoidi $e^{i2\pi ft}$ per tutte le frequenze f nel campo infinito $(-\infty, +\infty)$. La trasformata di Fourier risulta quindi come integrazione di componenti periodiche dalla quale risulta una funzione aperiodica.

Le sinusoidi $e^{i2\pi ft}$ hanno un'ampiezza infinitesimale

$$F(f) dt$$

cosicchè $F(f)$ non può essere interpretata come ampiezza caratteristica di $x(t)$ ma appare solo come grandezza relativa, esplicativa della distribuzione delle ampiezze su un campo continuo di frequenze. In questo caso non è più possibile parlare di funzione spettro ma di *funzione di densità spettrale*, definita come :

$$\Phi(f) = |F(f)|^2$$

Nel caso di funzioni casuali non si può procedere mediante l'usuale teoria delle serie e degli integrali di Fourier. La ragione è intuibile: quando si parla di funzioni casuali ci si riferisce a fenomeni in cui le fluttuazioni delle quantità osservate in funzione del tempo non possono essere pre

cisamente predette. Il metodo di analisi più indicato è quello di descrivere tali fenomeni attraverso strumenti analitici particolari, cioè tramite le funzioni di ripartizione. Ovviamente le funzioni di ripartizione non possono essere intese come espressioni matematiche deterministiche che leghino funzionalmente ed in modo univoco l'effetto del fenomeno considerato alla molteplicità delle cause che l'ha originato. In questo senso la ricomposizione della funzione $x(t)$ casuale originaria da funzioni descrittive, quali le funzioni di ripartizione, risulta assurda. Dovrebbe essere chiaro quindi che, se si interpreta il fenomeno come risultante dalla somma di un processo generatore di infinite armoniche, l'estrazione casuale delle medesime non fornisce univocamente una risposta ma una classe di possibili risposte ciascuna delle quali è vista come funzione membro appartenente ad essa e non può esistere nessuna operazione che corrisponde alla sintesi vista per le funzioni periodiche ed aperiodiche.

La risposta alla domanda se anche per le funzioni casuali esista una funzione dell'importanza analitica propria della funzione spettro diviene affermativa in seguito al teorema cosiddetto di Wiener-Khintchine.

Il teorema dimostra che una funzione spettro significativa ed utile per l'insieme di funzioni membro esiste ed è denominata *densità di spettro della funzione casuale o processo stocastico*. La costruzione concettuale su cui poggia il teorema supera la difficoltà relativa all'impossibilità (come detto in precedenza) di sintetizzare un processo stocastico mediante una particolare funzione membro, costruendo la densità di spettro tramite una funzione riassuntiva, sulla base del comportamento medio, delle informazioni da analizzare comuni ad ogni funzione membro.

Tale è la funzione di autocorrelazione definita come:

$$\varphi_{xx}(u) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t) x(t+u) dt$$

dove T è il periodo, u è lo sfasamento tra due istanti temporali. Si dimostra che la notazione soprariportata è la traduzione in termini "ergodici"

della scrittura più familiare:

$$\varphi_{1^x 2^x}^{1^x 2^x}(u) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{1^x 2^x}^{1^x 2^x}(x_1, x_2; u) dx_1 dx_2$$

dove $f_{1^x 2^x}^{1^x 2^x}$ è la funzione di frequenza congiunta delle v.c. 1^x e 2^x .

L'enunciazione del teorema di Wiener-Kintchine (si veda [12], pag.117-258 e [8], pag.604-615) è questa: la funzione di autocorrelazione di una funzione casuale e la funzione di densità di spettro della medesima sono in relazione tra loro attraverso la trasformata coseno di Fourier.

Il che porta a considerare le relazioni:

$$(5) \quad \varphi_{xx}(u) = \int_{-\infty}^{\infty} \phi_{xx}(f) \cos(2\pi fu) df$$

e

$$(6) \quad \phi_{xx}(f) = \int_{-\infty}^{\infty} \varphi_{xx}(u) \cos(2\pi fu) du$$

dove $\phi_{xx}(f)$ è la funzione di densità di spettro di una funzione casuale.

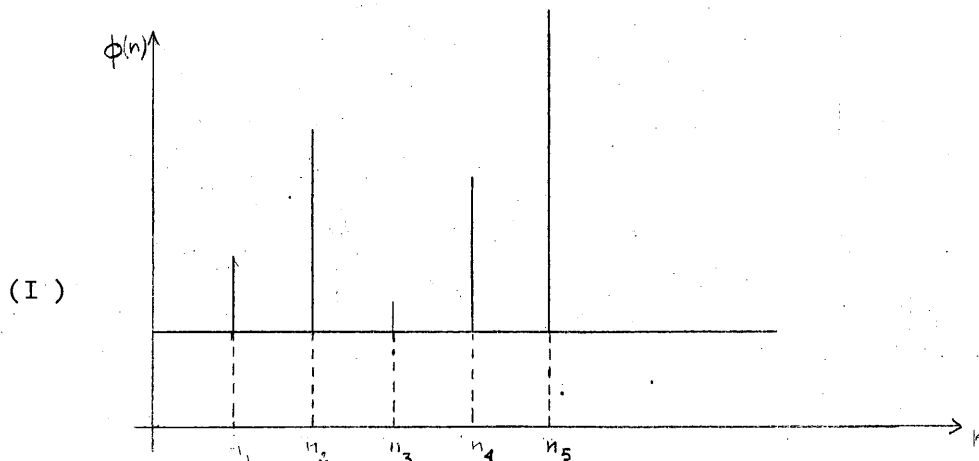
Dalla relazione (5) si deduce un'importante implicazione. Qualora si consideri la funzione

$$\xi(t) = x(t) - m \text{ dove } m = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t) dt, \text{ per } u = 0 \text{ la (5) diventa:}$$

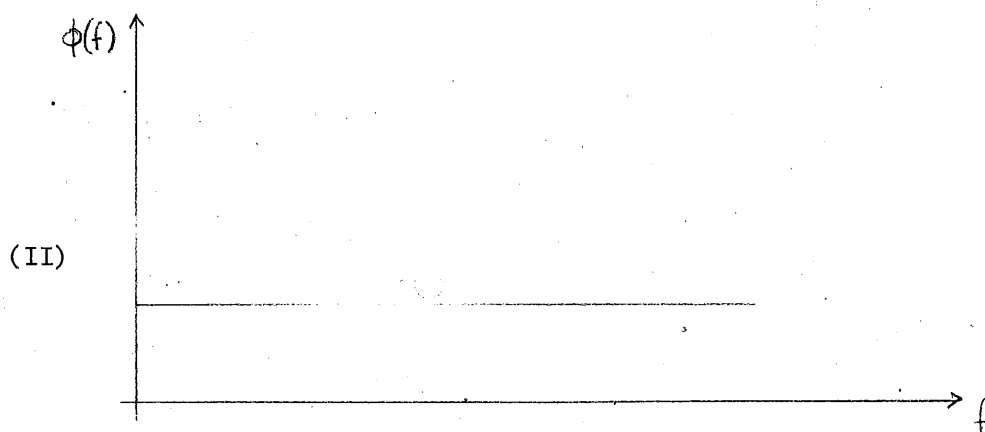
$$\varphi_{\xi\xi}^{(0)} = \sigma_{\xi\xi}^2 = \int_{-\infty}^{\infty} \phi_{\xi\xi}(f) df$$

Si evidenzia qui ancor più il significato della funzione di densità spettrale: esso spiega come la varianza del fenomeno si distribuisce sopra il dominio frequenziale. Rivedendo quanto affermato in precedenza per le funzioni periodiche, è opportuno sottolineare come la funzione risulti definita sopra

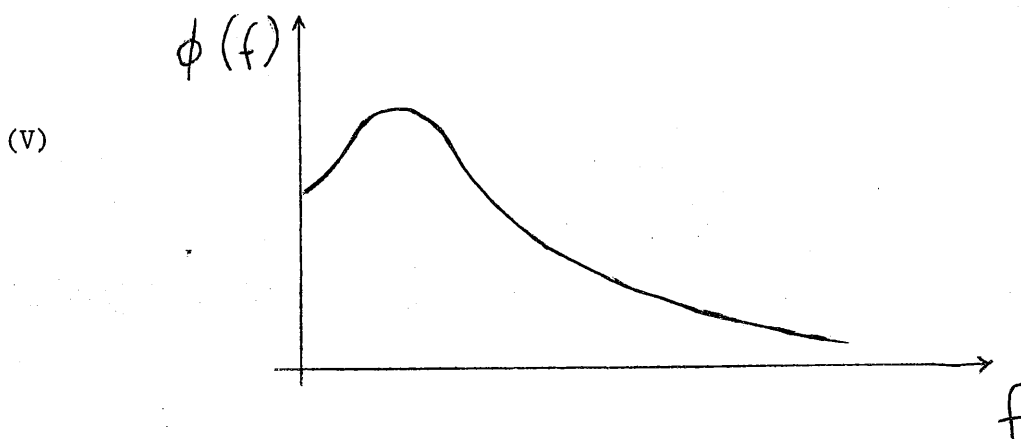
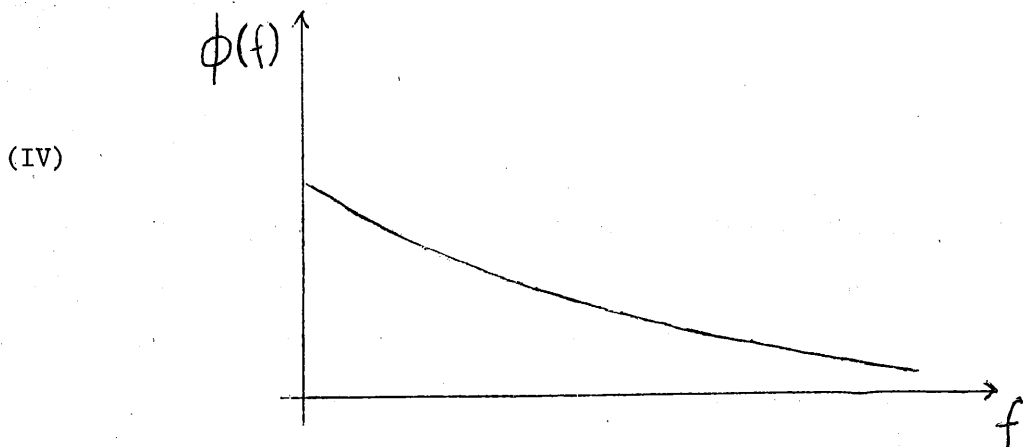
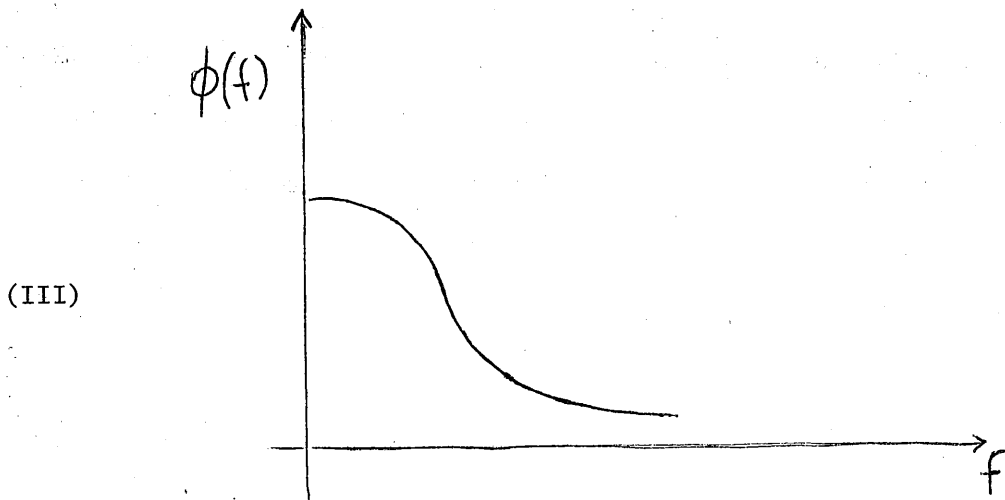
un dominio di frequenza discreto (per opportunità preso nel semiasse positivo). Graficamente si può definire uno spettro come un diagramma che indichi la misura delle ampiezze a ciascuna frequenza discreta cioè del tipo:



La funzione di densità spettrale di una funzione aperiodica o di una funzione casuale risultano definite sopra un dominio di frequenza continuo. Così per una funzione casuale i cui valori siano incorrelati tra loro (white noise), poichè nessuna frequenza contribuisce in misura maggiore o minore alla variabilità del fenomeno, presenterà, per una banda di frequenza abbastanza ampia, un diagramma del tipo:



Considerando, invece, un'onda semplicemente casuale priva di componenti deterministiche periodiche o aperiodiche si potrà rappresentare in un diagramma, per esempio, del tipo :



Il diagramma (III) rappresenta la densità spettrale del processo di Poisson cioè di un insieme di funzioni che assumono solo due valori k e $-k$ descrivendo così un cammino alternato in cui il numero dei passaggi tra le due costanti, cioè il numero dei cambiamenti di segno, entro un certo interval-

lo temporale definisce una variabile casuale con distribuzione poissoniana. I diagrammi (IV) e (V) pur attribuibili a processi di natura stocastica di vengono di difficile interpretazione. E' importante sottolineare che par-
tendo da modelli regressivi lineari si perviene a diagrammi spettrali di ta
li forme, ciò non significa però che diagrammi spettrali di questo tipo in-
dichino nel modello regressivo lineare il modello corretto: altri sistemi
molto più complicati possono portare a diagrammi della stessa forma (si ve
da /4_, pag.54).

Da quanto finora esposto risulta possibile definire una funzione $S(f)$ det-
ta funzione di spettro cumulato o spettro integrato che, essendo una funzio-
ne monotona non decrescente, potrà essere scritta nella classica scomposizio-
ne:

$$S(f) = S_1(f) + S_2(f) + S_3(f)$$

dove S_1 , S_2 ed S_3 sono funzioni non decrescenti delle frequenze. S_1 è
una funzione assolutamente continua con derivata $F(f)$; S_2 è una "step
function", ed S_3 è una funzione continua e costante quasi ovunque nel senso
matematico ed è chiamata *parte singola* della scomposizione. Di conseguenza,
si dimostra che qualsiasi processo stazionario può essere scomposto in
 $x_t = x_1(t) + x_2(t) + x_3(t)$ e qualora le 3 componenti additive siano incor-
relate, il loro spettro integrato risulta essere rispettivamente:

$$S_1(f), S_2(f), S_3(f)$$

La componente $x_2(t)$ viene denominata componente deterministica e corri-
sponde alla somma di un numero finito di termini armonici ed interpreta,
quindi, un processo lineare^(xx) ciclico; $x_1(t)$ è un processo appartenente
alla classe dei processi così detti non deterministici che contiene proces-
si lineari autoregressivi a media mobile e regressivi.

(xx) Per il concetto di "processo lineare" si veda /6_, pag.157

La componente $x_3(t)$ non è così facilmente caratterizzabile, può comunque essere dimostrato che è una componente deterministica, come si desume dalla natura ampiamente costante del suo spettro integrato. Nelle applicazioni economiche la componente corrispondente a $S_3(f)$ non sembra essere considerevolmente significativa e può essere assunta pari a zero.

Il significato della dizione "non deterministico" perviene dall'impossibilità, nel caso di $x_1(t)$, di predire esattamente il fenomeno. Qualora sia possibile determinare un previsore lineare H_t , ottimizzato con il metodo dei minimi quadrati, rimarrebbe sempre una varianza non nulla relativa al processo, cioè:

$$\epsilon_t = x_t - H_t$$

Se sul piano teorico è possibile definire una funzione spettrale o di densità spettrale e comprenderne il significato che la lega alla funzione $x(t)$ originaria, sul piano operativo non bisogna dimenticare che le serie temporali si presentano su un dominio temporale limitato. Ciò comporta l'introduzione dell'inferenza statistica. Autori come il Blackman - Tukey (in [1]) si sono posti il problema di applicare l'analisi teorica alle serie campionarie. Si è così arrivati alla determinazione della funzione Spettro Campionario e ad analizzare come si distribuisce la stima dello spettro di un processo stocastico del quale la serie storica considerata è vista come una delle determinazioni temporali casualmente possibili. Nella presente indagine ci si limiterà ad una analisi descrittiva delle serie esaminate senza fare nessuna assunzione sul tipo di distribuzione che le caratterizza e trascurando i problemi tipici dell'inferenza.

§.3 Problemi connessi alla non stazionarietà delle Serie Temporalì.

La classe più semplice di processi non stazionari, alla quale si farà riferimento nel presente lavoro, può essere interpretata dal modello:

$$y_t = m(t) + x_t$$

dove x_t è una serie stazionaria.

Non è difficile dimostrare (si veda [4], pag.129) che:

$$\text{Med} \{y_t\} = \bar{y} = \frac{1}{n} \sum_1^n m(t) + \text{Med}\{x_t\}$$

$$\text{Var} \{y_t\} = s_y^2 = \frac{1}{n} \sum_1^n m^2(t) - \frac{1}{n} \left[\sum_1^n m(t) \right]^2 + \frac{2}{n} \sum_1^n m(t) \cdot \frac{\bar{x}_t - \bar{x}}{n} + s_x^2$$

dove

$$s_x^2 = \frac{1}{n} \sum x_t^2 - \frac{1}{n} \bar{x}^2 ; \quad \bar{x} = \text{Med}\{x_t\}$$

Si può così arrivare a scindere la varianza del processo y_t nella componente dovuta al trend e nella componente dovuta alla variabilità del processo stazionario x_t , considerando:

$$\text{Med} \left\{ s_y^2 \right\} = \frac{1}{n} \sum m^2(t) - \frac{1}{n} \left[\sum_1^n m(t) \right]^2 + \sigma_x^2$$

La scomposizione nel dominio frequenziale della parte di varianza dovuta al trend, che rappresenta una funzione aperiodica definita in un intervallo finito, descriverà una funzione spettrale che assume un punto di massa di altezza proporzionale alla varianza nella frequenza zero e densità nulla altrove. Ciò è intuibile immaginando una funzione aperiodica come funzione periodica con periodo infinito.

La proprietà di linearità dei coefficienti di Fourier porterebbe a concludere che lo spettro del processo y_t è identico a quello del processo x_t eccetto che per l'ordinata relativa alla frequenza zero. In realtà, quando si deve operare con un campione effettivo, sorgono due diversi ordini di difficoltà: il primo riguarda la possibilità di distinguere il trend da andamenti ciclici con frequenza molto bassa, il secondo riguarda l'infiltrazione (leakage) di varianze tra bande di frequenza contigue e quindi una più difficile interpretazione della stima spettrale.

Riguardo alla prima difficoltà, quando si opera sulla serie effettiva di dati, diventa concettualmente impossibile distinguere quale tipo di funzione analitica ha determinato la concentrazione di massa spettrale nella frequenza zero. Ad esempio non sarà distinguibile un trend del tipo polinomiale:

$$m(t) = a_0 + \sum_{n=1}^h a_n t^n \quad t = 1, 2, \dots, N$$

a_n costanti reali

da un trend di tipo armonico, che assume la forma:

$$m(t) = a_0 + \sum_{n=1}^h a_n \cos(2\pi n f t + \theta_n) \quad t=1, 2, \dots, N \text{ ed } f < \frac{1}{N}$$

non esiste, cioè, nessun criterio analitico che ne stabilisca l'esatta natura. Qualora, però, si voglia porre l'attenzione sulle periodicità nascoste della serie relative a bande di frequenza diverse, sarà sufficiente trascurare la banda di frequenza attorno allo zero.

La possibilità di infiltrazione tra bande di frequenza è dovuta alla particolare lettura della stima spettrale, effettuata attraverso dei fattori di ponderazione parametrici rispetto alla frequenza; tali fattori di ponderazione possono essere pensati come una fessura o finestra (spectral window) che lascia filtrare certe oscillazioni armoniche su una certa fascia di frequenza, mentre sopprime o comprime le oscillazioni nelle fasce di frequenza diverse. E' proprio l'imperfezione analitica di tale finestra che permette l'infiltra-

zione di variabilità tra fasce di frequenza contigue.

Operare direttamente sul processo y_t porta, quindi, ad inevitabili errori di stima sulle ciclicità relative al processo x_t e questo tanto più è alta la variabilità di y_t dovuta alla componente trend. Di qui la necessità di usare le tecniche di detrendizzazione che portino alla determinazione del trend e alla sua eliminazione, tenendo conto dell'influenza che ciò può avere sulla funzione spettrale. La rimozione della componente trend può avvenire sia attraverso l'analisi regressiva sia con opportune tecniche di filtraggio; la sostanziale differenza tra i due metodi risiede nel fatto che l'uso dei filtri implica delle trasformazioni lineari della serie, mentre ciò non avviene con la regressione, ma non esistono regole generali di preferenza, caso per caso si dovrà scegliere la tecnica adatta. Il vantaggio nell'uso delle tecniche di regressione, per esempio, è dovuto al mantenimento della stessa numerosità dei dati originari, mentre ciò non succede per le tecniche di filtraggio che usano le differenze finite. Inoltre è stato dimostrato dal Blackman - Tukey (si veda [1], pag.139-146) che la regressione con una funzione polinomiale e la successiva rimozione non inficiano le stime spettrali al di fuori delle bande di frequenza zero dove la massa spettrale viene effettivamente rimossa.

§.4 Scopo dell'analisi

La critica che maggiormente viene indirizzata all'Indice Generale riguarda il sistema di ponderazione usato dall'ISTAT nella costruzione dell'indice. E' noto infatti come, al di là delle procedure di computo, il peso percentuale del capitolo alimentare superi notevolmente l'importanza dei rimanenti quattro capitoli di spesa. L'analisi spettrale, con la verifica della struttura delle serie temporali prese in esame, permette di rispondere ad alcuni interrogativi connessi alla scelta dei pesi operata dall'ISTAT: se tale scelta, cioè, riflette in modo inalterato la struttura degli andamenti ciclici dei singoli capitoli di spesa sull'indice generale e l'eventualità di compensazioni tra andamenti ciclici o di ciclicità indotte tra singoli capitoli. La possibilità di stabilire l'importanza di un capitolo di spesa piuttosto di un altro per le ciclicità indotte nell'indice generale non è legata solamente ad un problema di verifica, ma riguarda pure finalità di ordine operativo dal momento che il livello salariale nell'industria (e non solo in essa) segue l'andamento dell'indice del costo della vita attraverso il meccanismo della scala mobile. Il computo delle retribuzioni, infatti, viene effettuato tenendo conto di una indennità di contingenza variata di tre mesi in tre mesi e non può non risentire degli andamenti ciclici attorno alla linea del trend dell'indice generale, dato che gli scatti di contingenza o punti vengono calcolati in base a variazioni di cinquanta centesimi per gli scatti positivi e di 101 centesimi per gli scatti negativi^(x).

La media degli scarti assoluti attorno ai diversi trends calcolati varia per l'indice generale da un massimo di 116 centesimi per la retta ad un minimo di 66 centesimi per la cubica, così pure in riferimento ai singoli capitoli di spesa la media degli scarti assoluti si presenta leggermente superiore. Da questo punto di vista l'andamento oscillatorio dei capitoli di

(x)

Cfr. - Accordo del 15/1/1957 per la scelta mobile delle retribuzioni -
Supplemento ordinario alla G.U. n. 274 del 9/11/1960

spesa dovrebbero contribuire a spostare l'indice generale pur rimanendo de
terminante per il computo dei punti di contingenza l'andamento dei rispetti
vi trends.

§.5 Elaborazione dei dati e risultati

L'ISTAT costruisce l'indice del costo della vita in base 1938 = 1 per il periodo 1951 - 1963, in base 1961 = 100 per il periodo 1964 - 1966, in base 1966 = 100 per il periodo 1967 - 1970, in base 1970 = 100 dal 1971 in poi.

Poichè la metodologia spettrale rivela la sua utilità se applicata a serie stazionarie di sufficiente lunghezza campionaria, non avrebbero significato curve spettrali calcolate separatamente per le singole sottoserie; si è quindi affrontato il problema di omogeneizzare queste ultime sull'intero asse temporale. Per ragioni di semplificazione e per la scarsa numerosità dei dati con base 1970, si è tentato di costruire una serie omogenea e stazionaria per l'intero periodo dal 1951 al 1970. In tale ricostruzione non sono stati usati i coefficienti di raccordo delle pubblicazioni ufficiali data la loro arbitrarietà, preferendo detrendizzare le tre sottoserie (relative ai tre sottoperiodi con base diversa) per riunire in un unico vettore di numerosità pari a 240 le sottoserie dei residui. In questo modo eventuali alterazioni alle stime spettrali vengono ristrette alle sole bande di frequenza molto vicine allo zero. La mancanza di qualsiasi criterio analitico per la individuazione del trend ha suggerito l'opportunità di detrendizzare con un polinomio lineare nel tempo di primo, secondo e terzo grado e con una curva esponenziale. La stima delle curve spettrali calcolate sui quattro vettori dei residui relativa ad ognuna delle detrendizzazioni sono state corrette attraverso procedure già note in analisi spettrale come procedure di smussamento ed ottenute, come già accennato al §.3, ponderando le funzioni di autocorrelazione attraverso particolari fattori o finestre ritardo (lag window). Con tale procedimento si riesce in effetti a rendere più consistente le stime dello spettro anche se ciò va a scapito della loro correttezza. Quindi oltre alla "stima grezza" delle curve spettrali, si sono calcolate anche "stime smussate" corrispondenti alla scelta di due punti di troncamento nella finestra di PARZEN (si veda $\overline{7}$, pag.163-166) pari a 160 e 120 "lags" eliminando così le periodicità fittizie che apparivano nella curva grezza

dello spettro.

Per trovare un'ulteriore conferma ai risultati sono state calcolate anche le curve di coerenza fra gli spettri dell'indice generale e ciascun capitolo di spesa. Qualsiasi accenno a tale funzione è stato rimandato in questa sede per non appesantire la trattazione teorica esposta al §.2 ; vengo no qui ommesse le formulazioni (si veda /6 /) che necessiterebbero di una introduzione alla teoria degli spettri incrociati cercando di darne semplicemente il significato. La funzione di coerenza ("coherency" oppure anche "squared coherency") assume valori compresi tra zero ed uno e può essere interpretata in modo abbastanza simile al quadrato del coefficiente di correlazione per una variabile casuale bidimensionale. Può, infatti, indicare se esiste o no perfetta concordanza tra andamento ciclico di una serie e andamento ciclico della seconda, relativamente ad ogni frequenza. Per esempio, qualora per entrambe le serie vi sia, entro una determinata fascia di frequenza, un notevole contributo alla varianza del fenomeno da esse spiegato, la funzione di coerenza avrà, per la medesima fascia di frequenza, valori molto vicini ad uno. Con altro significato la funzione di coerenza indica quanta parte di densità (o massa) spettrale viene scomputata considerando il primo fenomeno come output di un processo lineare in cui il secondo viene considerato come input. I risultati ottenuti con il calcolo delle curve spettrali sono riportati nella tab. 2 . Accanto alle periodicità elencate sono riportate le percentuali di varianza spiegata dalla curva spettrale all'interno di fasce di frequenza che hanno per centro le periodicità stesse. Tali fasce sono costanti rispetto al dominio frequenziale e pari ad un intervallo di $3f$, ma variano rispetto al dominio temporale al variare della frequenza nell'intervallo $(0, \frac{1}{2})$. Infatti , indicando con f_n la frequenza espressa in cicli per unità di tempo, cioè:

$$f_n = \frac{n}{T} \quad n = 0, 1, 2, \dots, \frac{T}{2}$$

allora f_n^{-1} è determinata da una successione di punti che si addensano in

progressione armonica in prossimità di $f_{T/2}^{-1} = 2$. Il valore di $f_{T/2}^{-1}$ risulta due volte l'unità di misura temporale in cui viene espresso il fenomeno e costituisce l'intervallo minimo di osservazione del fenomeno allo interno del quale ha ancora senso parlare di ciclo. Per un periodo $T = 240$ la variazione temporale delle fasce di frequenza segue i valori indicati nella tab. 1 :

Tab. 1

Centro della Fascia		Fascia di Frequenza		Ampiezza Fascia Temporale
in frequenza (cicli per mese)	in unità di tempo = anno	(cicli per mese)		(in mesi)
1/3	0,25	$\frac{471}{1440}$	$\frac{489}{1440}$	0,31
1/6	0,5	$\frac{231}{1440}$	$\frac{249}{1440}$	0,5
1/12	1	$\frac{111}{1440}$	$\frac{129}{1440}$	1,7
1/18	1,5	$\frac{71}{1440}$	$\frac{89}{1440}$	3,8
1/24	2	$\frac{51}{1440}$	$\frac{69}{1440}$	6,8
1/30	2,5	$\frac{39}{1440}$	$\frac{57}{1440}$	12,8
1/36	3	$\frac{31}{1440}$	$\frac{49}{1440}$	15,6
1/48	4	$\frac{21}{1440}$	$\frac{39}{1440}$	26,5
1/60	5	$\frac{15}{1440}$	$\frac{33}{1440}$	46,4
1/72	6	$\frac{11}{1440}$	$\frac{29}{1440}$	70,4

Ai fini di una corretta interpretazione è necessario tener presente che le periodicità riportate pur essendo riferite al valore centrale della fascia (approssimandole al mese) e quindi esprimendo l'esatta periodicità a cui cor

Tab. 3 TABELLA DEI VALORI DI COERENZA TRA INDICE GENERALE
E SINGOLI CAPITOLI DI SPESA

PERIODICITA' anni mesi	ALIMENTARE	ABBIGLIAMENTO	ELETT. E COMB.	ABITAZIONE	SPESE VARIE
2 4	0,980	0,399	0,563	0,633	0,808
1 9	0,972	0,408	0,462	0,515	0,538
1 3	0,975	0,675	0,474	0,691	0,761
0 9	0,951	0,517	0,262	0,681	0,802

N.B. I valori di coerenza sono ottenuti su serie detrendizzate con la curva esponenziale.

risponde un'onda ciclica del tipo: $A_k \cos \left(\frac{2\pi kt}{T} + \theta_k \right)$, esse non trovano nessun significato economico, ma puramente matematico-statistico. Infatti si può presumibilmente escludere l'esistenza in campo economico di ciclicità perfettamente deterministiche come esistono invece nel campo della fisica. Il problema qui non è l'individuazione di una particolare ciclicità, ma semplicemente l'analisi strutturale delle serie attraverso la scomposizione della loro varianza nel dominio frequenziale e le relazioni che possono esistere tra la struttura di una serie e quelle di un'altra. Nella tab.3 sono indicati i risultati ottenuti con il calcolo delle curve di coerenza tra gli spettri dell'indice generale ed i singoli capitoli di spesa. Tali risultati sono espressi come valori medi dei valori di coerenza relativi alle stesse bande di frequenza per le quali è indicato il termine centrale dell'indice generale nella tab. 2. I risultati ottenuti dall'analisi dei dati possono essere così sintetizzati:

- 1) Perfetta analogia della curva spettrale relativa all'indice generale con la curva spettrale del capitolo alimentare.
- 2) Difformità tra gli spettri dei singoli capitoli di spesa.

L'analogia e la difformità sono immediatamente desumibili dall'esame delle ciclicità riportate nella tab. 2 che si ripetono identicamente per l'indice generale e per il capitolo alimentare con percentuale di varianza spiegata pressochè uguale, mentre nessuna delle ciclicità relative agli altri capitoli di spesa risulta riportata nella colonna dell'indice generale. L'unica eccezione potrebbe essere costituita dal capitolo spese varie, le cui ciclicità di 2 anni e 4 mesi e 0 anni e 9 mesi si confondono con le analoghe ciclicità del capitolo alimentare. L'analogia è ancor più immediata dalla lettura dei valori di coerenza, vicinissimi ad uno per il capitolo alimentare e relativamente bassi per gli altri capitoli di spesa.

Questo risultato è dovuto alle tecniche di computo e al tipo di ponderazione usato dall'ISTAT. Infatti il peso percentuale del capitolo alimentare sull'indice generale rispetto ai rimanenti quattro è stato del 66,40% per l'indice con base 1938, del 54,65% per l'indice con base 1961 del 48,61% per l'in

dice con base 1966 (si veda /5_7/).

Il significato dell'analogia riscontrata tra indice generale e capitolo alimentare è evidente: l'andamento ciclico del capitolo alimentare si riflette interamente sull'indice generale, mentre ciò non succede per gli altri capitoli di spesa.

Questi ultimi quindi non contribuiscono a spostare l'indice generale se non per la componente trend: circa le variazioni attorno alla linea del trend assumere l'indice generale equivale ad assumere il capitolo alimentare. Tale risultato potrebbe trovare una giustificazione economica qualora si considerino i prezzi dei beni al consumo alimentari come prezzi pilota, cioè come prezzi che inducono quelli degli altri beni di consumo. Un aumento dei prezzi delle derrate alimentari potrebbe provocare, attraverso l'aumento del costo della vita, incrementi salariali e quindi l'aumento dei prezzi anche per i settori dei beni di consumo non alimentari. Nelle indagini sulla situazione economica italiana si trovano dei modelli interpretativi dove il saggio di variazione dei salari è considerato variabile dipendente del saggio di variazione nel costo della vita (si veda /10_7/, pag.71) non esistono analisi sulla relazione più particolare tra il livello dei beni di consumo alimentare ed i beni di consumo non alimentare. Tuttavia è possibile escludere l'ipotesi dei prezzi dei beni di consumo alimentare come prezzi guida per gli altri beni di consumo: sia perchè ciò contraddirebbe il risultato di difformità tra gli spettri dei singoli capitoli di spesa sia perchè i prezzi dei beni di consumo alimentare sono per lo più determinati dai prezzi dei prodotti agricoli mentre i beni di consumo non alimentare sono per lo più determinati dai prezzi dei prodotti industriali destinati al consumo. E' risaputo (si veda /10_7/, pag.53) come nel primo caso il mercato dei prodotti agricoli rispecchi ancora condizioni di concorrenza e quindi i prezzi nel breve periodo dipendono dalla domanda ed offerta, mentre per i prezzi dei prodotti industriali prevalgono condizioni oligopolistiche e di conseguenza, nel breve periodo, essi dipendono dall'andamento dei costi. In ultima analisi, un aumento dei prezzi dei prodotti industriali potrebbe essere indotto da incrementi sala-

riali, ma da una parte tali incrementi non dipendono solo dal saggio di variazione del costo della vita bensì dalla percentuale di disoccupazione e dall'azione sindacale, dall'altra l'incremento dei costi può aversi per cause diverse dagli incrementi salariali, per esempio, come aumento del costo dei fattori. Ciò spiega la possibile difformità ottenuta con l'analisi empirica dei dati.

La lettura degli spettri ha confortato l'assunto teorico evidenziando la sostanziale indifferenza nell'uso di un trend relativamente alle fasce di frequenza lontane dalla frequenza zero. Per le frequenze relativamente vicine, quindi per le periodicità più alte, si è avuta una loro leggera modificazione, ne sono esempio per il capitolo dell'abbigliamento il passaggio dai 3 anni e 10 mesi relativi alla detrendizzazione con la retta e con la curva esponenziale ai circa 3 anni per la parabola e la cubica, e per l'elettricità e combustibili il passaggio dai 3 anni e 6 mesi (retta ed esponenziale) ai 2 anni e 3 mesi circa (parabola e cubica).

Un'altra importante osservazione che lega il calcolo degli spettri al tipo di detrendizzazione riguarda il capitolo dell'abitazione. L'importanza percentuale della ciclicità annua riscontrabile nella detrendizzazione con la parabola, la cubica ed anche con la esponenziale non trova riscontro nella detrendizzazione con la retta. Ciò è immediatamente interpretabile con l'inadeguatezza della retta a spiegare l'evoluzione del fenomeno nel tempo, cioè nell'incapacità della retta di sottrarre completamente alla totalità del fenomeno quella parte di varianza attribuibile alla componente trend e quindi nell'infiltrazione di tale varianza, non assorbita completamente nella fascia di frequenza zero, in altre bande di frequenza. Il fenomeno dell'infiltrazione è pure desumibile dal confronto delle percentuali tra le diverse detrendizzazioni: periodicità più basse vengono esaltate se si è considerato un trend di grado più elevato, il contrario ovviamente succede per le periodicità più alte.

§.6 Ricostruzione dell'indice generale del costo della vita detrendizzato

I risultati dell'analisi condotta sui singoli capitoli di spesa ha suggerito l'opportunità di ricostruire l'indice generale detrendizzato con il metodo delle componenti principali e verificare una maggiore sensibilità di tale metodo nel sussumere le ciclicità di tutti i cinque capitoli di spesa. La determinazione delle componenti principali si rivela assai utile in questo caso poichè, oltre ad esprimere il maggior contributo possibile alla variabilità complessiva delle variabili originarie, conduce ad una combinazione lineare di queste che, come il Girshick (si veda [3], pag.519) ha dimostrato, minimizza la varianza imputabile alle eventuali irregolarità accidentali e mantiene la massima correlazione possibile con ciascuna variabile originaria.

Nel calcolo delle componenti principali si è usato il metodo diretto determinando i coefficienti dagli autovettori dalla matrice di varianze e covarianze o dalla matrice di correlazione mantenendo le quattro diverse detrendizzazioni.

L'operare con numeri indici, riferiti ad una medesima base, poteva permettere la formulazione dell'ipotesi di uguale ordine di grandezza delle variabili e quindi la determinazione degli autovettori dalla matrice di varianze e covarianze.

Essendo però necessario, ai fini di un confronto tra curve spettrali, ottenere per la prima componente principale un'area della massa spettrale pari a quella dell'indice generale data dall'ISTAT, si è ricostruito l'indice detrendizzato in modo che il suo ordine di grandezza non fosse riferito alle singole variabili originarie, cioè ai singoli capitoli di spesa, ma a tale indice generale.

Si è quindi proceduto alla standardizzazione delle variabili originarie calcolando le prime componenti principali in base agli autovettori della matrice di correlazione per poi ridimensionarle formulando l'ipotesi che, pur prescindendo dalla struttura della serie ricostruita, la sua varianza coincidesse con la varianza dell'indice generale detrendizzato ISTAT.

Indicando con ${}^i\xi_t$ le variabili standardizzate relative ai cinque capitoli di spesa, ($i=1,2,\dots,h=5$; $t=1,2,\dots,N=240$), con ${}^i\eta_t$ le componenti principali standardizzate e con λ_1 l'autovalore di modulo maggiore, sono note le seguenti relazioni:

$${}^j\xi_t = \sum_{i=1}^h c_{ij} {}^i\eta_t$$

$$c_{ij} = \frac{a_{ij} \sqrt{\lambda_1}}{\sqrt{\sum_{i=1}^h a_{ij}^2}}$$

dove (a_{ij}) è la matrice degli autovettori della matrice di correlazione

$${}^i\eta_t = \sum_{j=1}^h {}^{ij}c {}^j\xi_t$$

dove $({}^{ij}c) = (c_{ij})^{-1}$

In particolare :

$${}^1\eta_t = \sum_{j=1}^h {}^{1j}c {}^j\xi_t$$

Volendo esprimere la variabile ${}^1\eta_t$ come funzione lineare delle variabili originarie sarà sufficiente considerare:

$${}^j\xi_t = \frac{j_{X_t} - j_{\mu_x}}{j_{\sigma_x}}$$

ed essendo j_{X_t} variabili detrendizzate, sarà $j_{\mu=0}$, quindi :

$${}^1\eta_t = \sum_{j=1}^h \frac{{}^{1j}c}{j_{\sigma_x}} j_{X_t}$$

Anche ${}^1\eta_t$ potrebbe essere ridotto in termini di variabile dimensionata 1Y qualora fossero noti ${}^1\mu_y$ e ${}^1\sigma_y$.

Essendo ${}^1\eta_t$ la standardizzazione di una variabile supposta detrendizzata sarà ${}^1\mu_y = 0$ e ${}^1\sigma_y$ verrà posto uguale allo scarto quadratico medio del-

l'indice generale detrendizzato dell'ISTAT ottenendo così il medesimo ordine di grandezza in relazione alla variabilità lasciandone inalterata la diversa struttura. Inoltre è facilmente dimostrabile che porre tale uguaglianza equivale a minimizzare la somma dei quadrati degli scarti tra indice generale detrendizzato ISTAT e prima componente principale. I risultati delle elaborazioni sono riportati qui di seguito, relativamente ad ogni determinazione, mettendo in evidenza il contributo percentuale della prima componente principale sul totale della varianza:

$$\begin{aligned} \text{RETTA} \quad {}^1\eta_t &= 0,212845 \quad {}^1\xi_t + 0,341706 \quad {}^2\xi_t + 0,111379 \quad {}^3\xi_t + 0,284506 \quad {}^4\xi_t + \\ (51,50) \quad &+ 0,364647 \quad {}^5\xi_t \end{aligned}$$

$$\begin{aligned} \text{PARABOLA} \quad {}^1\eta_t &= 0,332967 \quad {}^1\xi_t + 0,145105 \quad {}^2\xi_t + 0,316355 \quad {}^3\xi_t - 0,221294 \quad {}^4\xi_t + \\ (50,15) \quad &+ 0,343327 \quad {}^5\xi_t \end{aligned}$$

$$\begin{aligned} \text{CUBICA} \quad {}^1\eta_t &= 0,357614 \quad {}^1\xi_t + 0,372909 \quad {}^2\xi_t + 0,27668 \quad {}^3\xi_t - 0,090357 \quad {}^4\xi_t + \\ (41,98) \quad &+ 0,3517361 \quad {}^5\xi_t \end{aligned}$$

$$\begin{aligned} \text{ESPONENZIALE} \quad {}^1\eta_t &= 0,295338 \quad {}^1\xi_t + 0,243133 \quad {}^2\xi_t + 0,215694 \quad {}^3\xi_t - 0,275545 \quad {}^4\xi_t + \\ (51,64) \quad &+ 0,344219 \quad {}^5\xi_t \end{aligned}$$

dove ${}^j\xi_t$ indica le variabili standardizzate relative ai numeri indici dei singoli capitoli di spesa e precisamente, nell'ordine, del capitolo Alimentare, Abbigliamento, Elettricità e Combustibili, Abitazione e del capitolo Spese Varie.

La relativamente modesta percentuale di varianza spiegata dalla prima componente principale è interpretabile con la scarsa correlazione delle serie stazionarie tra loro, quindi nella mancanza di un andamento ciclico struttu

ralmente analogo, come avevano già evidenziato le curve spettrali. Anche qui nel calcolo della prima componente principale si può osservare l'ineguatezza del modello di detrendizzazione rappresentato dalla retta.

Infatti l'incapacità della retta di assorbire completamente l'andamento evolutivo nel capitolo delle abitazioni altera le correlazioni di questo ultimo e gli altri capitoli di spesa determinando, a differenza delle altre detrendizzazioni, la positività del coefficiente di ξ_t^4 .

La presenza di un coefficiente negativo, per questo capitolo, propone un problema interpretativo in termini di struttura ciclica. E' quanto verrà affrontato nel prossimo paragrafo.

§.7 Alcune considerazioni tra analisi di Fourier e metodo delle componenti principali.

L'analisi di Fourier applicata alle serie temporali stazionarie bivariate^(x) può mettere in evidenza le relazioni che sussistono tra coefficienti di Fourier della serie e funzione di correlazione incrociata. Infatti, date due serie temporali, stazionarie e periodiche, esprimibili in serie di Fourier :

$$f_1(t) = \frac{1}{2} \frac{a_0}{1} + \sum_{n=1}^{\infty} (\frac{a_n}{1} \cos 2 \pi nft + \frac{b_n}{1} \sin 2 \pi nft)$$

$$f_2(t) = \frac{2}{2} \frac{a_0}{2} + \sum_{n=1}^{\infty} (\frac{a_n}{2} \cos 2 \pi nft + \frac{b_n}{2} \sin 2 \pi nft)$$

si può costruire la funzione di correlazione incrociata:

$$\varphi_{12}(t) = \sum_{n=-\infty}^{\infty} \bar{A}_1(n) A_2(n) e^{i 2 \pi nft}$$

dove

$$A_1(n) = \frac{1}{2} \frac{a_n - i b_n}{1} \quad \text{per } n = 0, \pm 1, \dots$$

$$A_2(n) = \frac{2}{2} \frac{a_n - i b_n}{2} \quad \text{per } n = 0, \pm 1, \dots$$

Ora ponendo

$$\alpha_n = \sqrt{a_n^2 + b_n^2}$$

$$\theta_n = \arctang \left(- \frac{b_n}{a_n} \right)$$

(x) Con il termine serie temporali bivariate si intende due serie di n dati quantitativi poste in corrispondenza mediante il tempo o più brevemente ogni coppia di serie definite sopra un'asse temporale (bivariate time series - si veda [6] pag.321)

allora

$$A(n) = \frac{\alpha}{2} e^{i\theta_n n}$$

Conoscendo le relazioni tra i coefficienti di Fourier

$$a_{-n} = a_n ; b_{-n} = -b_n \text{ e quindi } \theta_{-n} = -\theta_n$$

la funzione di correlazione può essere espressa come:

$$(7) \quad \varphi_{12}(\tau) = \sum_{-\infty}^{\infty} \frac{1^{\alpha}_n 2^{\alpha}_n}{4} \cos (2 \pi f n \tau + \theta_{2n} - \theta_{1n})$$

Analogamente può essere ricavata la funzione:

$$(8) \quad \varphi_{21}(\tau) = \sum_{-\infty}^{\infty} \frac{1^{\alpha}_n 2^{\alpha}_n}{4} \cos (2 \pi f n \tau + \theta_{1n} - \theta_{2n})$$

Il prodotto dei coefficienti armonici, nella struttura della $\varphi(\tau)$, indica che è sufficiente la mancanza di una armonica in una delle due serie temporali perchè sia assente anche nella correlazione incrociata. Inoltre dalla (7) e (8) è possibile ricavare la relazione:

$$\varphi_{12}(-\tau) = \varphi_{21}(\tau)$$

Nel presente lavoro, la necessità di operare con più di due serie temporali porta alla considerazione della matrice delle funzioni di correlazione incrociata, che indicheremo con $V = (\varphi_{ij}(\tau))$.

La matrice V non è simmetrica rispetto alla diagonale principale, essa è però riducibile alla ben nota matrice di varianze e covarianze qualora le funzioni di correlazione incrociata siano considerate solo nel punto $\tau = 0$. Infatti si può osservare come per $\tau = 0$

$$\varphi_{21}(0) = \varphi_{12}(0)$$

essendo:

$$\varphi_{12}^{(0)} = \sum_{-\infty}^{\infty} \frac{1_n^{\alpha} 2_n^{\alpha}}{4} \cos (2_n^{\theta} - 1_n^{\theta})$$

$$\varphi_{21}^{(0)} = \sum_{-\infty}^{\infty} \frac{1_n^{\alpha} 2_n^{\alpha}}{4} \cos (1_n^{\theta} - 2_n^{\theta})$$

ed essendo la funzione coseno una funzione pari.

Qualora si consideri il passaggio dalla matrice di varianza e covarianza alla matrice di correlazione

$$R = \left(\frac{\varphi_{ij}^{(0)}}{\varphi_{ii}^{(0)} \varphi_{jj}^{(0)}} \right)$$

muta l'ordine di grandezza delle ampiezze delle armoniche ma viene lasciata inalterata la loro fase, come è possibile desumerlo esplicitando il rapporto:

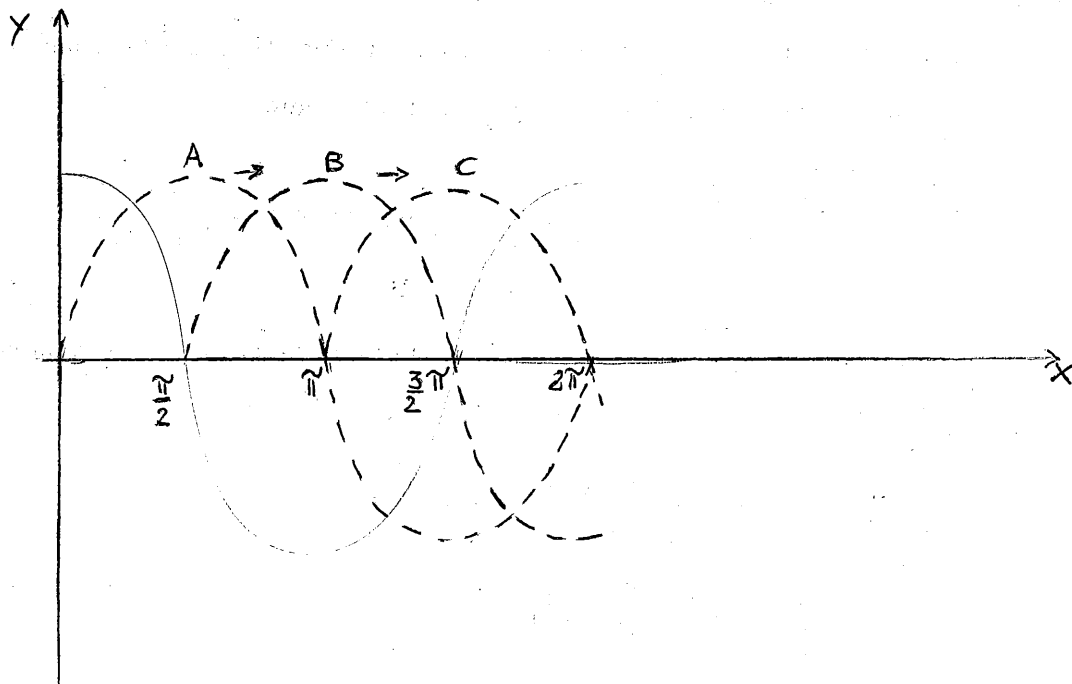
$$r_{ij} = \frac{\varphi_{ij}^{(0)}}{\sqrt{\varphi_{ii}^{(0)} \varphi_{jj}^{(0)}}} = \frac{\sum_{-\infty}^{\infty} i_n^{\alpha} j_n^{\alpha} \cos (j_n^{\theta} - i_n^{\theta})}{\sqrt{\sum_{-\infty}^{\infty} i_n^{\alpha 2} \sum_{-\infty}^{\infty} j_n^{\alpha 2}}}$$

La struttura dei coefficienti di correlazione indica che il segno, positivo o negativo, dei coefficienti stessi dipende dalla differenza di fase tra le serie correlate. Si consideri il caso particolare in cui le serie temporali sono composte di una sola armonica. Si avrà:

$$r_{ij} = \frac{\varphi_{ij}^{(0)}}{\sqrt{\varphi_{ii}^{(0)} \varphi_{jj}^{(0)}}} = \frac{i^{\alpha} j^{\alpha} \cos (j^{\theta} - i^{\theta})}{\sqrt{i^{\alpha 2} j^{\alpha 2}}} = \cos (j^{\theta} - i^{\theta})$$

E' facile comprendere come, in questo caso, la condizione per avere un coefficiente di correlazione negativo risulti essere una differenza di fase compresa nell'intervallo $(\frac{\pi}{2}, \frac{3}{2} \pi)$. L'andamento delle sinusoidi che rispet

tano tale condizione può essere visualizzata dal seguente grafico:



Nel grafico sono state riportate solo tre curve sfasate rispetto alla cosinusoide $y = \cos x$, cioè:

A) $y = \cos \left(x + \frac{\pi}{2} \right)$

B) $y = \cos (x + \pi)$

C) $y = \cos \left(x + \frac{3}{2}\pi \right)$

A e C rappresentano le curve relative agli estremi dell'intervallo di negatività del coefficiente di correlazione, mentre B rappresenta la curva la cui correlazione con la cosinusoide $y = \cos x$ è negativa e massima, cioè è uguale a -1 .

Nel caso più generale di coefficienti di correlazione strutturalmente rappresentati da una somma infinita di termini, relativi cioè a serie temporali scomponibili in un numero infinito di armoniche, l'interpretazione della negatività del coefficiente di correlazione è più complicata. Dato però che la funzione $\varphi_{ij}(z)$ contiene solo le armoniche presenti in entrambe le serie

temporali, è possibile dare un'interpretazione analoga a quella vista nel caso della singola armonica.

Considerando infatti che l'espressione della funzione di correlazione può essere riscritta, essendo una funzione pari, come segue:

$$\varphi_{ij}(\tau) = \frac{i^a_0 j^a_0}{4} + \frac{1}{2} \sum_1^{\infty} i^{\alpha}_n j^{\alpha}_n \cos(2\pi n \tau + j^{\theta}_n - i^{\theta}_n)$$

e ponendo i valori medi i^a_0 e j^a_0 delle rispettive serie pari a zero, il coefficiente r può così formularsi:

$$r_{ij} = \frac{\sum_1^{\infty} i^{\alpha}_n j^{\alpha}_n \cos(j^{\theta}_n - i^{\theta}_n)}{\sqrt{\sum_1^{\infty} i^{\alpha 2}_n \sum_1^{\infty} j^{\alpha 2}_n}}$$

Inoltre è possibile approssimare r trascurando i valori di i^{α}_n e j^{α}_n molto prossimi allo zero calcolando:

$$r_{ij} \approx \frac{\sum_1^N i^{\alpha}_n j^{\alpha}_n \cos(j^{\theta}_n - i^{\theta}_n)}{\sqrt{\sum_1^N i^{\alpha 2}_n \sum_1^N j^{\alpha 2}_n}}$$

Ciò equivale a determinare i coefficienti di Fourier approssimando con il metodo dei minimi quadrati le serie temporali, supposte continue nell'intervallo $(-T/2, T/2)$, mediante le somme parziali:

$$\frac{a_0}{2} + \sum_1^N (a_n \cos 2\pi n f t + b_n \sin 2\pi n f t)$$

E' proprietà delle approssimazioni ai minimi quadrati con funzioni ortogonali determinare ciascun coefficiente in modo indipendente da tutti gli altri e dal numero di armoniche impiegate nell'approssimazione.

L'operazione di troncamento, effettuata al grado di approssimazione desiderato, non altera quindi l'ordine di grandezza dei coefficienti α_n e non influisce sul segno di r . Affinchè sia $r < 0$ sarà sufficiente che le N armoniche presenti in entrambe le serie temporali siano sfasate nell'intervallo $\left(\frac{\pi}{2}, \frac{3}{2}\pi\right)$. In ogni caso sarà necessario che il contributo all'andamento delle serie determinato dalle armoniche sfasate nell'intervallo $\left(\frac{\pi}{2}, \frac{3}{2}\pi\right)$ prevalga complessivamente sul contributo delle altre armoniche. Per esigenze di sinteticità d'espressione diremo "in antifase" siffatte serie.

La premessa, fin qui svolta, sul significato da attribuire alla negatività (o positività) del coefficiente di correlazione induce alla verifica del legame tra i segni degli elementi della matrice di correlazione ed i segni dei coefficienti l_j relativi alla ricostruzione della prima componente principale l_t .

Occorre subito dire che tale legame si riconduce a quello esistente tra matrice di correlazione e autovettore soluzione del sistema lineare omogeneo costruito con l'autovalore di modulo maggiore λ_1 . Infatti si indichi con C la matrice dei coefficienti c_{ij} , si potrà scrivere:

$$C = (\Lambda)^{\frac{1}{2}} A$$

dove A è la matrice degli autovettori a_j ($j=1,2,\dots,h$) normalizzati ^(x) e Λ la matrice diagonale degli autovalori.

Essendo A una matrice ortogonale, cioè uguale alla sua matrice contragrediente ^(xx) A^{-1} , risulta:

(x) Si ricorda che un vettore X' è normalizzato rispetto al vettore X se la sua norma è $\|X'\| = \sqrt{{}^t(X')X'} = 1$

(xx) Se P è una matrice invertibile, l'uguaglianza $PP^{-1} = I$ permette di scrivere ${}^t(P^{-1})P = {}^tI = I$ dimostrando che una trasposta di una matrice invertibile è anch'essa invertibile ed è uguale alla trasposta della matrice P^{-1} , detta matrice contragrediente

$$C^{-1} = \left[(\Lambda)^{\frac{1}{2}} A \right]^{-1} = {}^t \left[(\Lambda)^{\frac{1}{2}} A \right] = (\Lambda)^{\frac{1}{2}} {}^t A$$

Quindi C^{-1} mantiene gli stessi segni di ${}^t A$ e poichè si può scrivere A nella forma:

$$A = \begin{pmatrix} a_j \\ -j \\ \|a_j\| \end{pmatrix} \quad (j=1, 2, \dots, h)$$

il fatto che sia, per definizione, $\|a_j\| \geq 0$ permette di stabilire la seguente relazione rispetto al segno positivo o negativo degli elementi A_{ij} di A e gli elementi a_{ij}

$$\text{sgn}(A_{ji}) = \text{sgn}(a_{ji}) \quad \forall_i, \forall_j .$$

E' possibile, a questo punto, dare una condizione sufficiente sulla matrice R affinchè si abbia un solo elemento negativo dell'autovettore a_1 . Si ha infatti il seguente teorema:

Se la matrice simmetrica \bar{P} (p_1, p_2, \dots, p_n) possiede un solo vettore $\bar{p}_j < 0$ e quindi simmetricamente ${}^t(\bar{p}_j) < 0$ eccetto l'elemento $p_{jj} > 0$, allora l'autovettore $'z = ('z_1, 'z_2, \dots, 'z_h)$ corrispondente all'autovalore di modulo massimo di \bar{P} possiede uno e un solo elemento $'z_j < 0$ ($j=1, 2, \dots, h$)

Dimostrazione:

Per dimostrare tale teorema è opportuno rifarsi al lemma di Perron, dimostrato in [2], relativo ad una proprietà degli autovalori ed autovettori di una qualsiasi matrice a elementi reali positivi detta matrice positiva. Esso dice: una matrice positiva P , di ordine h , possiede sempre un autovalore reale e positivo λ che è una radice semplice dell'equazione caratteristica ed è superiore ai moduli di tutti gli altri autovalori. A questo autovalore "massimale" λ , corrisponde un autovettore $z = (z_1, z_2, \dots, z_h)$ di P di coordinate positive $z_i > 0$ ($i=1, 2, \dots, h$).

La matrice simmetrica \bar{P} può essere pensata come una matrice P positi-

va con il vettore $p_j > 0$ moltiplicato per lo scalare $k = -1$; si dovrà quindi dimostrare che:

a) detto $p_1^{\lambda_1}$ l'autovalore di modulo massimo di P e $\bar{p}_1^{\lambda_1}$ l'autovalore di modulo massimo di \bar{P} risulta:

$$\bar{p}_1^{\lambda_1} = p_1^{\lambda_1}$$

ed in conseguenza a tale uguaglianza:

b) detto (z_1, z_2, \dots, z_h) l'autovettore relativo a $(P - p_1^{\lambda_1} I)$ risulta:

$$\begin{cases} z_i = z_i & \text{per } i \neq j \\ z_i = -z_i & \text{per } i = j \end{cases}$$

Essendo la matrice \bar{P} simmetrica dovrà essere:

$$\bar{p}_j = {}^t(\bar{p}_j)$$

quindi se:

$$(9) \quad \bar{p}_j = (-1) p_j$$

risulta

$$(10) \quad {}^t(\bar{p}_j) = (-1) {}^t(p_j)$$

Ciò equivale alla moltiplicazione, nella matrice P , dell' j -esimo vettore colonna per lo scalare $k_1 = -1$ e dell' j -esimo vettore riga per lo scalare $k_2 = -1$ cosicchè rimane immutato l'elemento p_{jj} e, valendo per i de-terminanti la proprietà:

$$(11) \quad k |P| = |(p_1, p_2, \dots, kp_j, \dots, p_h)|$$

rimane inalterato anche il polinomio caratteristico

$$\Delta(\lambda) = |P - \lambda I| = |p_{ij} - \lambda \delta_{ij}|$$

dove δ_{ij} è il simbolo di Kronecker.

Resta allora dimostrato quanto affermato in a) .

Riguardo a b) si osservi che gli elementi z_i ($i=1,2,\dots,h$) sono i minori di ordine $h-1$ estraibili dalla matrice $(P - \lambda_1 I)$ sopprimendo il vettore riga linearmente dipendente e l' i -esimo vettore colonna. In base alle relazioni (9), (10) e (11) già viste risulterà allora:

$$\left\{ \begin{array}{ll} z_i = z_i k_1 k_2 & \text{per } i \neq j \\ z_i = z_i k_1 & \text{per } i = j \end{array} \right.$$

Poichè la matrice di correlazione R soddisfa i requisiti del teorema una possibile interpretazione sulla presenza di un coefficiente negativo l_j^c e quindi di un elemento negativo dell'autovettore si potrà esprimere affermando che tutte le serie considerate risultano in antifase rispetto alla j -esima serie.

Nel presente lavoro tale condizione è verificata nelle matrici di correlazione relative alle serie detrendizzate con la parabola e con l'esponenziale.

§.8 Analisi spettrale del ricostruito indice generale del costo della vita detrendizzato.

La verifica della capacità della prima componente principale ad assumere le caratteristiche della struttura ciclica dei cinque capitoli di spesa è stata effettuata operando con l'analisi spettrale sulle quattro prime componenti principali relative alle quattro detrendizzazioni, ottenendo così i valori riportati nella tab.4, analoga alla tab.1 .

Dai risultati ottenuti è possibile formulare alcune osservazioni:

- a) Rispetto all'indice Generale ISTAT detrendizzato, l'indice Generale ricostruito riesce ad assumere una nuova ciclicità di bassa frequenza con periodo attorno ai 3 anni e 10 mesi risentendo quindi di analoghe basse frequenze presenti nei capitoli dell'Abbigliamento e dell'Elettricità e Combustibili
- b) Sono completamente assenti le ciclicità di più alta frequenza relative al capitolo dell'Abitazione come era facilmente prevedibile dato che la relativa serie risulta in antifase (nel senso definito al paragrafo precedente) con tutte le altre e quindi viene compensato il suo contributo ciclico dalle medesime ciclicità sfasate presenti negli altri capitoli di spesa. In corrispondenza alla periodicità di 1 anno, 6 mesi e 3 mesi osservate nel capitolo dell'abitazione (si veda tab.2) la struttura dello spettro delle quattro prime componenti presentano dei punti di minimo.
- c) L'effetto della detrendizzazione operata per i singoli capitoli di spesa non si riflette in modo omogeneo sulla ricostruzione dell'Indice Generale attraverso il metodo delle componenti principali:
all'aumentare del grado del polinomio interpolante la massa spettrale rimossa vicino alla banda di frequenza zero ha un comportamento anomalo .
Ciò induce a ipotizzare la possibilità di alterazioni delle ciclicità, indotte all'uso del metodo delle componenti principali, la cui verifica esce, comunque, dalle finalità impostesi al presente lavoro.

Tab. 4 Indice Generale ricostruito del Costo della Vita

RETTA			PARABOLA			CUBICA			ESPONENZIALE		
PERIODO	% di	Var.	PERIODO	% di	Var.	PERIODO	% di	Var.	PERIODO	% di	Var.
anni	mesi		anni	mesi		anni	mesi		anni	mesi	
3	10	21,50	3	10	15,52	3	10	18,16	3	10	15,76
2	4	9,07	2	4	9,80	2	4	12,83	2	4	14,31
1	9	3,96	1	9	8,99	1	9	4,85	1	9	15,52
1	3	2,42	1	3	0,81	1	3	2,94	1	3	5,58
0	11	1,11	0	11	0,54	0	11	0,45	0	11	2,15
0	9	0,96	0	9	0,98	0	9	1,84	0	9	1,19
0	8	0,93	0	8	0,62	0	8	1,07	0	8	0,45

L'uso del metodo delle componenti principali si è, quindi, rivelato maggiormente sensibile alla ricomposizione, nella prima componente principale, delle ciclicità presenti nei singoli capitoli di spesa; si può sottolineare come esso permetta, attraverso il segno negativo dei coefficienti, l'evidenziarsi di ciclicità "in antifase" tra singole serie, cioè la possibilità di compensazione tra andamenti oscillatori opposti nel loro andamento. Tuttavia va fatto notare, come questo metodo, già usato per la costruzione dello Indice Generale completo della sua componente di Trend, parta dal presupposto di inalterabilità tra correlazioni, cioè porti alla costruzione di un modello in cui la concordanza tra capitoli di spesa non sia alterabile da mese a mese. L'impossibilità di mantenere valido questo presupposto anche nel breve periodo - si pensi, per esempio, all'influenza di una qualsiasi variazione dei prezzi relativi al consumo sui coefficienti di correlazione - pone delle perplessità sull'utilizzazione di tale metodo.

In ogni caso, si ricorda che nel presente lavoro si è resa possibile la sua utilizzazione per serie relativamente ampie nel tempo dopo aver separatamente detrendizzato le serie suddivise nei tre sottoperiodi ciascuno relativo ad una diversa base. L'applicazione di tale metodo ai singoli capitoli di spesa non detrendizzati, oltre a presentare l'inconveniente della diversa base, potrebbe anche rivelarsi di scarsa utilità riguardo alla ciclicità qualora le correlazioni, nel riflettere l'andamento dei trends, non riescano ad assorbire la parte di variabilità dovuta all'andamento oscillatorio attorno ad essi.

BIBLIOGRAFIA

- /1/ R.B. BLACKMAN AND J.W. TUKEY, *The measurement of Power Spectra*.
Dover, Publications, Inc., N.Y., (1958) .
- /2/ F.R. GANTMACHER, *Theorie des matrices*. Tome 2, Dunod, Paris, (1966).
- /3/ M.A. GIRSHICK, Principal components. In *J. Amer. Statist. Assoc.*,
n. 195 , (1936) .
- /4/ C.W.J. GRANGER and M. HATANAKA, *Spectral Analysis of Economic Time
Series*. Princeton University Press, (1964) .
- /5/ ISTAT, Numeri indici dei prezzi, base 1966 = 100 . In *Metodi e Norme-
Serie A* , n. 6 , (1967) .
- /6/ G.M. JENKINS and G.D. WATTS, *Spectral Analysis and its applications*.
Holden-Day, (1968).
- /7/ G.M. JENKINS, General Considerations in the Analysis of Spectra. In
Technometrics, 3, (1961) .
- /8/ A.Y. KHINTCHINE, Korrelations theorie der Stationaren stochastischen
Prozesse. In *Math. Ann.*, n. 109, (1934) .
- /9/ W.Y. LEE, *Statistical theory of communication*. John Wiley F. Sons ,
Inc., N.Y., (1960) .
- /10/ P. SYLOS LABINI, *Sindacati, inflazione e produttività*. Laterza, Bari
(1972) .
- /11/ R.C. TITCHMARSCH, *The theory of Functions*. Oxford University Press ,
(1932) .
- /12/ N. WIENER, Generalised harmonic analysis. In "*Acta Math*", n. 55,
(1930) .