

Handbook of Statistics

VOLUME 3

Time Series in the Frequency Domain

Edited by
D. R. Brillinger

P. R. Krishnaiah

Time Series in the Frequency Domain

Edited by

D. R. Brillinger

Department of Statistics
University of California, Berkeley, CA, U.S.A.

P. R. Krishnaiah

Center for Multivariate Analysis
University of Pittsburgh, Pittsburgh, PA, U.S.A.

1983

NORTH-HOLLAND
AMSTERDAM · NEW YORK · OXFORD

Preface

Time series analysis is one of the most flourishing of the fields of present day statistics. Exciting developments are taking place: in pure theory and in practice, with broad relevance and with narrow intent, for large samples and for small samples. The flourishing results in part, from the dramatic increase in the availability of computing power for both number crunching and for graphical display and in part from a compounding of knowledge as more and more researchers involve themselves with the problems of the field.

This volume of the *Handbook of Statistics* is concerned particularly with the frequency side, or spectrum, approach to time series analysis. This approach involves essential use of sinusoids and bands of (angular) frequency, with Fourier transforms playing an important role. A principal activity is thinking of systems, their inputs, outputs, and behavior in sinusoidal terms. In many cases, the frequency side approach turns out to be simpler in each of computational, mathematical, and statistical respects. In the frequency approach, an assumption of stationarity is commonly made. However, the essential roles played by the techniques of complex demodulation and seasonal adjustment show that stationarity is far from a necessary condition. So too are assumptions of Gaussianity and linearity commonly made. As various of the papers in this Volume show, nor are these necessary assumptions.

The Volume is meant to represent the frequency approach to time series analysis as it is today. Readers working their way through the papers and references included will find themselves abreast of much of contemporary spectrum analysis.

We wish to express our deep appreciation to Professors E. J. Hannan and M. B. Priestley for serving as members of the editorial board. Thanks are due to Professors P. Guttorp, E. J. Hannan, T. Hasan, J. Liljestøl, and M. B. Priestley for refereeing various chapters in the volume. We are most grateful to the authors and North-Holland Publishing Company for their excellent cooperation in bringing out this volume.

D. R. Brillinger
P. R. Krishnaiah

Contributors

- R. J. Bhansali, *Statistics Department, University of Liverpool, P.O. Box 147, Liverpool L69 3BX, U.K.* (Ch. 1)
- David R. Brillinger, *Department of Statistics, University of California, Berkeley, California 94720* (Ch. 2)
- T. C. Chang, *Department of Mathematical Sciences, University of Cincinnati, Cincinnati, Ohio 45221* (Ch. 20)
- William S. Cleveland, *Bell Laboratories, Murray Hill, New Jersey 07974* (Ch. 3)
- Robert B. Davies, *Department of Scientific and Industrial Research, Applied Mathematics Division, P.O. Box 1335, Wellington, New Zealand* (Ch. 4)
- Robert Engle, *Department of Economics, University of California, San Diego, La Jolla, California 92093* (Ch. 5)
- C. W. J. Granger, *Department of Economics, University of California, San Diego, La Jolla, California 92093* (Ch. 5)
- E. J. Hannan, *Department of Statistics, The Australian National University, Box 4, G.P.O., Canberra, A.C.T., Australia 2600* (Ch. 6)
- T. Hasan, *Bell Laboratories, Murray Hill, New Jersey 07974* (Ch. 7)
- Melvin J. Hinich, *Department of Government, University of Texas at Austin, Austin, Texas 78712* (Ch. 8)
- D. Karavellas, *Statistics Department, University of Liverpool, P.O. Box 147, Liverpool L69 3BX, U.K.* (Ch. 1)
- L. H. Koopmans, *Department of Mathematics, University of New Mexico, Albuquerque, New Mexico 87131* (Ch. 9)
- P. R. Krishnaiah, *Center for Multivariate Analysis, University of Pittsburgh, Pittsburgh, Pennsylvania 15260* (Ch. 20)
- J. C. Lee, *Bell Laboratories, Murray Hill, New Jersey 07974* (Ch. 20)
- R. Douglas Martin, *Department of Statistics, University of Washington, Seattle, Washington 98195* (Ch. 10)
- Emanuel Parzen, *Institute of Statistics, Texas A & M University, College Station, Texas 77843* (Ch. 11)
- J. Pemberton, *Department of Mathematics, University of Salford, Salford, M5 4WT, U.K.* (Ch. 12)

- M. B. Priestley, *Department of Mathematics, University of Manchester Institute of Science & Technology, P.O. Box 88, Manchester, M60 1QD, U.K.* (Ch. 13)
- T. Subba Rao, *Department of Mathematics, University of Manchester Institute of Science & Technology, P.O. Box 88, Manchester, M60 1QD, U.K.* (Ch. 14)
- Enders A. Robinson, *Department of Geophysics, University of Tulsa, Tulsa, Oklahoma 74104* (Ch. 15)
- P. M. Robinson, *Department of Mathematics, University of Surrey, Guildford, Surrey, GU2 5XH, U.K.* (Ch. 16)
- M. Rosenblatt, *Department of Mathematics, University of California, San Diego, La Jolla, California 92093* (Ch. 17)
- R. H. Shumway, *Division of Statistics, University of California, Davis, California 95616* (Ch. 18)
- Tony Thrall, *Systems Applications, Inc., 101 Lucas Valley Rd., San Rafael, California 94903* (Ch. 19)
- H. Tong, *Department of Statistics, Chinese University of Hong Kong, New Territories, Hong Kong* (Ch. 12)

Table of Contents

Preface v

Contributors xiii

Ch. 1. Wiener Filtering (with emphasis on frequency-domain approaches) 1 *R. J. Bhansali and D. Karavellas*

1. Introduction 1
2. Derivation of the filter transfer function and the filter coefficients 2
3. Realization of the Wiener filter in some special cases 7
4. Estimation of the Wiener filter 9
5. Asymptotic properties of the estimated filter coefficients 12
6. A simulation study 15
- References 18

Ch. 2. The Finite Fourier Transform of a Stationary Process 21 *D. R. Brillinger*

1. Introduction 21
2. Central limit theorems 26
3. Direct estimation of second-order spectra 28
4. Linear models 29
5. Parametric models 32
6. Other topics 33
- Appendix 35
- References 36

Ch. 3. Seasonal and Calendar Adjustment 39 *W. S. Cleveland*

1. Introduction 39
2. Methods for seasonal adjustment 47

- 3. Calendar adjustment 55
- 4. Graphics for seasonal and calendar adjustment 60
- 5. Spectrum analysis 65
- References 70

Ch. 4. Optimal Inference in the Frequency Domain 73

R. B. Davies

- 1. Introduction 73
- 2. Theory of asymptotic inference 74
- 3. Inference in the frequency domain 82
- 4. Summary and example 89
- References 91

Ch. 5. Applications of Spectral Analysis in Econometrics 93

C. W. J. Granger and R. Engle

- 1. Beginnings 93
- 2. Applications of the power spectrum 94
- 3. Application of the cross spectrum 96
- 4. Evaluation of seasonal adjustment procedures 99
- 5. Spectral regression 100
- 6. Frequency-domain factor analysis 103
- 7. Advanced techniques 104
- 8. Conclusions 106
- References 106

Ch. 6. Signal Estimation 111

E. J. Hannan

- 1. Introduction 111
- 2. Regression problems 113
- 3. Delay estimation 116
- Bibliographic Notes 122
- References 123

Ch. 7. Complex Demodulation: Some Theory and Applications 125

T. Hasan

- 1. Introduction 125
- 2. Basics of complex demodulation 126
- 3. Statistical properties 132
- 4. A class of amplitude modulated cosinusoids 148
- References 155

Ch. 8. Estimating the Gain of A Linear Filter from Noisy Data 157
M. J. Hinich

Introduction 157

1. Linear filters 157
2. Estimating the phase and gain of the transfer function 158
3. Use of the Hilbert transform 161
4. Numerical approximations 162
5. Estimating the log gain 164
- Appendix 166
- References 167

Ch. 9. A Spectral Analysis Primer 169
L. H. Koopmans

1. Introduction 169
2. A historical perspective 169
3. The time-series model 171
4. Spectral representations 173
5. The different types of spectra 175
6. Spectra and linear filters 175
7. Spectral parameters for bivariate time series 179
8. Spectra of multidimensional processes 181
- References 183

Ch. 10. Robust-Resistant Spectral Analysis 185
R. D. Martin

1. Introduction 185
2. Time-series outliers and nonnormality 187
3. Autoregressive spectral density estimates 196
4. Smoothed periodogram estimates 199
5. Robustness-resistance properties of the spectrum estimate 209
6. An example 212
7. Concluding remarks 217
- References 217

Ch. 11. Autoregressive Spectral Estimation 221
E. Parzen

Introduction 221

1. Correlation and spectral functions of a stationary time series 222
2. Filter models of a short-memory stationary time series 225
3. Entropy interpretation of autoregressive spectral densities 228

- 4. Parametrizations of autoregressive spectral estimators 233
- 5. Empirical autoregressive spectral estimation 234
- 6. Autoregressive order determination 238
- 7. Suggestions for empirical spectral analysis 239
- 8. A bibliography of autoregressive spectral estimation 242
- References 243

Ch. 12. Threshold Autoregression and Some Frequency-Domain Characteristics 249

J. Pemberton and H. Tong

- 1. Introduction 249
- 2. Some motivation 250
- 3. A simplest TAR model 252
- 4. Some generalisation and some frequency-domain characteristics 255
- 5. Diagnostic checks of SETAR models from frequency-domain viewpoint 266
- References 272

Ch. 13. The Frequency-Domain Approach to the Analysis of Closed-Loop Systems 275

M. B. Priestley

- 1. Introduction 275
- 2. Identification of linear systems 276
- 3. Closed-loop systems: Basic assumptions and notation 278
- 4. Conditions on the transfer function 280
- 5. Estimation of the system transfer function (A) 280
- 6. The case of linear feedback plus noise 286
- 7. Numerical estimation of the system transfer function and noise structure 287
- Summary 290
- References 290

Ch. 14. The Bispectral Analysis of Nonlinear Stationary Time Series with Reference to Bilinear Time-Series Models 293

T. Subba Rao

- 1. Introduction 293
- 2. Stationary time series and second-order spectral analysis 293
- 3. Second-order covariance and spectral analysis of linear time-series models 294
- 4. Bilinear time-series models and second-order spectral analysis 297
- 5. Volterra series expansion of $VBL(p)$ model 301
- 6. Spectral and bispectral density functions and their estimation 303
- 7. Bispectral analysis of the bilinear time-series model $BL(1, 0, 1, 1)$ 308
- 8. Tests for departure from linearity and Gaussianity 310
- Conclusions 318
- References 318

Ch. 15. Frequency-Domain Analysis of Multidimensional Time-Series Data 321

E. A. Robinson

1. Introduction 321
2. The wave equation 322
3. Frequency-domain analysis 323
4. The distributed source model in seismic exploration 327
5. Migration of seismic data 330
6. One-dimensional migration 331
7. Two-dimensional migration 335
8. Two-dimensional example 338
9. Concluding remarks 342
- References 342

Ch. 16. Review of Various Approaches to Power Spectrum Estimation 343

P. M. Robinson

1. Introduction 343
2. Quadratic spectrum estimators 350
3. Nonquadratic spectrum estimators 356
4. Spectrum estimation from incomplete data 359
- References 365

Ch. 17. Cumulants and Cumulant Spectra 369

M. Rosenblatt

- Introduction 369
1. Definition 369
2. Spectra 371
3. Spectral estimates 375
4. Nonlinear systems 378
5. Non-Gaussian linear processes and phase information 379
6. Applications 381
- References 381

Ch. 18. Replicated Time-Series Regression: An Approach to Signal Estimation and Detection 383

R. H. Shumway

1. Introduction 383
2. Estimation of the regression functions 386
3. Tests of hypotheses in the frequency domain 390
4. Applications to signal detection 396
5. Discussion 405
- References 406

Ch. 19. Computer Programming of Spectrum Estimation 409
T. Thrall

1. Introduction 409
2. Two methods of estimation 411
3. Fast Fourier transform (FFT) algorithms 417
4. Tapering and padding the data 419
5. Variance approximations 423
6. Bandwidths and degrees of freedom 425
7. Prefiltering and recoloring 428
8. Trend removal 431
9. Missing values 432
10. Other problems in spectrum estimation and computation 435
- References 436

Ch. 20. Likelihood Ratio Tests on Covariance Matrices and Mean Vectors of Complex Multivariate Normal Populations and their Applications in Time Series 439
P. R. Krishnaiah, J. C. Lee and T. C. Chang

1. Introduction 439
2. Complex multivariate normal and complex Wishart distributions 440
3. Distribution of the determinant of the complex multivariate beta matrix 441
4. Test for independence of sets of variates 442
5. Test for sphericity 444
6. Test specifying the covariance matrix 445
7. Test for multiple homogeneity of the covariance matrices 445
8. Simultaneous tests for the homogeneity of populations 447
9. Test specifying the values of the covariance matrix and mean vector 449
10. Applications in time series in the frequency domain 450
- Appendix 456
- References 475

Subject Index 477

Contents of Previous Volumes 483

Wiener Filtering (with emphasis on frequency-domain approaches)

R. J. Bhansali and D. Karavellas

1. Introduction

Let $\{y_t, x_t\}$ ($t = 0, \pm 1, \dots$) be a bivariate process. An important class of problems considered in time-series analysis may be formulated in terms of the problem: How can we best predict y_t from $\{x_s, s \leq t\}$? If $y_t = x_{t+\nu}$, $\nu > 0$, then the problem is that of predicting the 'future' of x_t on the basis of its past. If $x_t = \xi_t + \zeta_t$, where ζ_t is 'noise' and ξ_t the 'signal' and $y_t = \xi_{t+\nu}$, then for $\nu = 0$ the problem is that of 'signal extraction', for $\nu > 0$ that of predicting the signal and for $\nu < 0$ that of interpolating the signal, in the presence of noise. If y_t and x_t are arbitrary, then the problem is simply that of predicting one series from another. This last problem is itself of interest in a number of disciplines: for example, in Economics, interest is often centred on obtaining a distributed lag relationship between two economic variables (see, e.g., Dhrymes [11]) such as level of unemployment and the rate of inflation.

A complete solution to the problem of predicting y_t from the past, $\{x_s, s \leq t\}$, of x_t would consist of giving the conditional probability distribution of the random variable y_t when the observed values of the random variables $\{x_s, s \leq t\}$ are given. However, this is seldom practicable as finding such a conditional distribution is usually a formidable problem. A simplifying procedure of taking the mean value of this conditional distribution as the predictor of y_t is also rarely feasible because this mean value is in general a very complicated function of the past x 's. Progress may, however, be made if $\{y_t, x_t\}$ is assumed to be jointly stationary and attention is restricted to the consideration of the linear least-squares predictor of y_t , i.e. the best predictor, \hat{y}_t , say, of y_t is chosen from the comparatively narrow class of linear functions of $\{x_s, s \leq t\}$,

$$\hat{y}_t = \sum_{j=0}^{\infty} h(j)x_{t-j}, \quad (1.1)$$

the coefficients $h(j)$ being chosen on the criterion that the mean square error of prediction

$$\eta^2 = E(\hat{y}_t - y_t)^2 \quad (1.2)$$

be a minimum.

Formation of \hat{y}_t from the $\{x_s, s \leq t\}$ may be viewed as a filtering operation applied to the past of x_t and, especially in engineering literature, \hat{y}_t is known as the Wiener filter.

It should be noted that if $\{y_t, x_t\}$ is Gaussian, then the linear least-squares predictor, \hat{y}_t , of y_t is also the best possible predictor in the sense that it minimises the mean square error of prediction within the class of all possible predictors of y_t ; hence for the Gaussian case the consideration of only linear predictors is not a restriction.

2. Derivation of the filter transfer function and the filter coefficients

Suppose that $\{y_t, x_t\}$ ($t = 0, \pm 1, \dots$) is real-valued jointly stationary with zero means, i.e. $Ex_t = Ey_t = 0$. If the means are nonzero, then these may be subtracted out. Let $R_{xx}(u) = E(x_{t+u}x_t)$ and $R_{yy}(u) = E(y_{t+u}y_t)$ denote the autocovariance functions of x_t and y_t , respectively, and let $R_{yx}(u) = E y_{t+u}x_t$ denote their cross-covariance function. Assume that

$$\sum_{u=-\infty}^{\infty} |R_{xx}(u)| < \infty, \quad \sum_{u=-\infty}^{\infty} |R_{yy}(u)| < \infty, \quad \sum_{u=-\infty}^{\infty} |R_{yx}(u)| < \infty$$

and let

$$f_{xx}(\lambda) = (2\pi)^{-1} \sum_{u=-\infty}^{\infty} R_{xx}(u) \exp(-iu\lambda),$$

$$f_{yy}(\lambda) = (2\pi)^{-1} \sum_{u=-\infty}^{\infty} R_{yy}(u) \exp(-iu\lambda)$$

denote the power spectral density functions of x_t and y_t , respectively, and

$$f_{yx}(\lambda) = (2\pi)^{-1} \sum_{u=-\infty}^{\infty} R_{yx}(u) \exp(-iu\lambda)$$

their cross-spectral density function. Assume also that $f_{xx}(\lambda) \neq 0$ ($-\infty < \lambda < \infty$).

Under these conditions x_t has the one-sided moving average representation (see Billinger [9, p. 78])

$$x_t = \sum_{j=0}^{\infty} b(j) \varepsilon_{t-j}, \quad b(0) = 1, \quad (2.1)$$

and the autoregressive representation

$$\sum_{j=0}^{\infty} a(j)x_{t-j} = \varepsilon_t, \quad a(0) = 1. \quad (2.2)$$

Here ε_t is a sequence of uncorrelated random variables with 0 mean and finite variance σ^2 , say, and the $\{b(j)\}$ and $\{a(j)\}$ are absolutely summable coefficients, i.e. they satisfy

$$\sum_{j=0}^{\infty} |b(j)| < \infty, \quad \sum_{j=0}^{\infty} |a(j)| < \infty.$$

Also, if

$$B(z) = \sum_{j=0}^{\infty} b(j)z^j, \quad A(z) = \sum_{j=0}^{\infty} a(j)z^j, \quad (2.3)$$

respectively, denote the characteristic polynomials of the $b(j)$ and the $a(j)$, then $B(z) \neq 0$, $A(z) \neq 0$, $|z| \leq 1$ and $A(z) = \{B(z)\}^{-1}$. The transfer functions $B(e^{-i\lambda})$ and $A(e^{-i\lambda})$ of the $b(j)$ and $a(j)$ are denoted by $B(\lambda)$ and $A(\lambda)$ respectively. We have $A(\lambda) = \{B(\lambda)\}^{-1}$ and $f_{xx}(\lambda) = \sigma^2(2\pi)^{-1}|B(\lambda)|^2$.

If $f_{xx}(\lambda)$ is known exactly, then the $\{b(j)\}$ and $\{a(j)\}$ may be determined, by the Wiener-Hopf spectral factorization procedure (Wiener [25, p. 78]). The assumptions made previously on $R_{xx}(u)$ and $f_{xx}(\lambda)$ ensure that $\log f_{xx}(\lambda)$ is integrable and hence has the Fourier series expansion

$$\log f_{xx}(\lambda) = \sum_{v=-\infty}^{\infty} c(v) \exp(-iv\lambda), \quad (2.4)$$

with

$$c(v) = (2\pi)^{-1} \int_{-\pi}^{\pi} \log f_{xx}(\lambda) \exp(iv\lambda) d\lambda \quad (2.5)$$

and

$$\sum_{v=-\infty}^{\infty} |c(v)| < \infty.$$

Set

$$B(\lambda) = \exp \left\{ \sum_{v=1}^{\infty} c(v) \exp(-iv\lambda) \right\}, \quad (2.6)$$

$$A(\lambda) = \{B(\lambda)\}^{-1} \quad (2.7)$$

and

$$\sigma^2 = 2\pi \exp\{c(0)\}. \quad (2.8)$$

Then

$$b(j) = (2\pi)^{-1} \int_{-\pi}^{\pi} B(\lambda) \exp(ij\lambda) d\lambda, \quad (2.9)$$

$$a(j) = (2\pi)^{-1} \int_{-\pi}^{\pi} A(\lambda) \exp(ij\lambda) d\lambda, \quad (2.10)$$

and the $\{b(j)\}$ and $\{a(j)\}$ thus obtained are absolutely summable (Brillinger [9,

p. 79]); see also Doob [12, pp. 160–164] and Grenander and Rosenblatt [16, pp. 67–81] for related work.

Next, consider prediction of y_t from the past, $\{x_s, s \leq t\}$, of x_t and in particular the determination of the filter coefficients $h(j)$ of the linear least-squares predictor \hat{y}_t of y_t . The mean square error of prediction η^2 is given by

$$\eta^2 = R_{yy}(0) - 2 \sum_{j=0}^{\infty} h(j) R_{yx}(j) + \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} h(j) h(k) R_{xx}(k-j). \quad (2.11)$$

If the $h(j)$ minimise η^2 , then we must have $\partial \eta^2 / \partial h(j) = 0$ ($j = 0, 1, \dots$). This requirement leads to the equations

$$\sum_{k=0}^{\infty} h(k) R_{xx}(k-j) = R_{yx}(j) \quad (j = 0, 1, \dots). \quad (2.12)$$

That the $h(k)$ satisfying (2.12) also minimise η^2 may be established by using an argument analogous to that given, for example, by Jenkins and Watts [18, pp. 204–205].

Equations (2.12) provide discrete analogues of the Wiener–Hopf integral equations (Wiener [25, p. 84]). As their left-hand side is of the form of a convolution, the use of Fourier series techniques is a natural approach to adopt for solving them. However, as discussed by N. Levinson (see [25, p. 153]) a direct use of the Fourier series techniques for obtaining the $h(j)$ is not feasible as well, because (2.12) is valid only for $j \geq 0$. Therefore, a somewhat indirect approach is adopted for expressing $h(j)$ in terms of $f_{yx}(\lambda)$ and $f_{xx}(\lambda)$.

The representation (2.1) implies that

$$R_{xx}(u) = \sigma^2 \sum_{s=0}^{\infty} b(s) b(s+u) \quad (u = 0, 1, \dots). \quad (2.13)$$

Put

$$D(\lambda) = f_{yx}(\lambda) \overline{A(\lambda)} = \sum_{u=-\infty}^{\infty} d(u) e^{-iu\lambda}, \quad (2.14)$$

and

$$[D(\lambda)]_+ = \sum_{u=0}^{\infty} d(u) \exp(-iu\lambda), \quad (2.15)$$

where

$$d(u) = (2\pi)^{-1} \int_{-\pi}^{\pi} f_{yx}(\lambda) \overline{A(\lambda)} \exp(iu\lambda) d\lambda \quad (2.16)$$

and

$$\sum_{u=-\infty}^{\infty} |d(u)| < \infty.$$

Note that $2\pi d(u) = E(y_t \varepsilon_{t-u})$ and $D(\lambda)$ gives the cross-spectral density function of y_t and ε_t .

From (2.14), we get

$$R_{yx}(j) = 2\pi \sum_{s=0}^{\infty} b(s)d(j+s). \quad (2.17)$$

Hence, (2.12) may be rewritten as

$$\sum_{s=0}^{\infty} b(s)d(j+s) = \frac{\sigma^2}{2\pi} \sum_{s=0}^{\infty} \sum_{k=0}^{\infty} h(k)b(s)b(s+j-k) \quad (j=0, 1, \dots),$$

or, as

$$d(v) = \frac{\sigma^2}{2\pi} \sum_{k=0}^{\infty} h(k)b(v-k) \quad (v=0, 1, \dots). \quad (2.18)$$

Since, $b(v) = 0$, $v < 0$, (2.18) may be solved by the Fourier series techniques. On multiplying both the sides of (2.18) by $e^{-i\lambda v}$ and summing for all $v \geq 0$, we get

$$\begin{aligned} H(\lambda) &= \sum_{k=0}^{\infty} h(k) \exp(-ik\lambda) \\ &= \frac{2\pi}{\sigma^2} B(\lambda)^{-1} [D(\lambda)]_+ = \frac{2\pi}{\sigma^2} A(\lambda) [f_{yx}(\lambda) \overline{A(\lambda)}]_+, \end{aligned} \quad (2.19)$$

and

$$h(j) = (2\pi)^{-1} \int_{-\pi}^{\pi} H(\lambda) \exp(ij\lambda) d\lambda. \quad (2.20)$$

Since the $d(u)$ given by (2.16) and the $a(j)$ given by (2.9) are absolutely summable, so are the $h(j)$ (see, e.g., Fuller [14, p. 120]). Thus, the $h(j)$'s satisfy

$$\sum_{j=0}^{\infty} |h(j)| < \infty.$$

The mean square error of prediction η^2 is

$$\begin{aligned} \eta^2 &= E\{(y_t - \hat{y}_t)^2\} = R_{yy}(0) - \sum_{j=0}^{\infty} h(j)R_{yx}(j) \\ &= \int_{-\pi}^{\pi} \left[f_{yy}(\lambda) - \left\{ \frac{2\pi}{\sigma^2} \|[D(\lambda)]_+\|^2 \right\} \right] d\lambda = R_{yy}(0) - \frac{4\pi^2}{\sigma^2} \sum_{j=0}^{\infty} d^2(j). \end{aligned} \quad (2.21)$$

Equations (2.19) and (2.21) are consistent with the results of Whittle [24, pp. 66-68], but note that a dividing factor of σ^2 is missing in equation (3.7.2) of Whittle [24, p. 42]; see also Bhansali [3].

It is instructive to compare the 'one-sided' predictor (2.10) with the corresponding 'two-sided' predictor of y_t obtained by assuming that the complete past, present and the complete future of x_t is known. Let

$$\bar{y}_t = \sum_{j=-\infty}^{\infty} g(j)x_{t-j} \quad (2.22)$$

be the 'two-sided' linear least-squares predictor of y_t . Then, as in (2.12), the $g(j)$ are the solutions of the equations

$$R_{yx}(u) = \sum_{j=-\infty}^{\infty} g(j)R_{xx}(u-j) \quad (u = 0, \pm 1, \dots). \quad (2.23)$$

Since these equations are two-sided and are valid for all integral values of u , they may be solved by the Fourier series techniques. On multiplying both the sides of (2.23) by $(2\pi)^{-1} e^{-iu\lambda}$ and summing over u , we have

$$\Gamma(\lambda) = \sum_{j=-\infty}^{\infty} g(j) \exp(-ij\lambda) = f_{yx}(\lambda)/f_{xx}(\lambda) \quad (2.24)$$

and

$$g(j) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \Gamma(\lambda) \exp(ij\lambda) d\lambda. \quad (2.25)$$

Let $\tau^2 = E\{(y_t - \bar{y}_t)^2\}$ be the corresponding mean square error of prediction. We have

$$\tau^2 = \int_{-\pi}^{\pi} \left\{ f_{yy}(\lambda) - \frac{|f_{yx}(\lambda)|^2}{f_{xx}(\lambda)} \right\} d\lambda = \int_{-\pi}^{\pi} \{1 - C_{yx}^2(\lambda)\} f_{yy}(\lambda) d\lambda, \quad (2.26)$$

where $C_{yx}(\lambda) = |f_{yx}(\lambda)| / \{f_{yy}(\lambda)f_{xx}(\lambda)\}^{1/2}$ is called the coherence between y_t and x_t . Note that $0 \leq C_{yx}(\lambda) \leq 1$, all λ . Expression (2.26) therefore shows that if $C_{yx}(\lambda)$ is close to 1 at all frequencies, then τ^2 is close to 0, and one would expect to obtain a close linear fit between y_t and x_t . In this sense, $C_{yx}(\lambda)$ may be interpreted as a correlation coefficient 'in the frequency domain' (see, e.g., Priestley [22] and Granger and Hatanaka [15]).

On using (2.14)–(2.16), (2.21) may be rewritten as (see Whittle [24, p. 69])

$$\eta^2 = \tau^2 + \frac{4\pi^2}{\sigma^2} \sum_{j=-\infty}^{-1} d^2(j). \quad (2.27)$$

The second term to the right of this expression gives the increase in mean square error due to the restriction that only the 'past' of x_t may be used for predicting y_t . In general, therefore, $\eta^2 \geq \tau^2$.

There is, however, one important situation in which $\eta^2 = \tau^2$. This occurs when x_t is the input to, and y_t the output of, a physically realizable linear time-invariant filter with uncorrelated noise, i.e. when,

$$y_t = \sum_{j=0}^{\infty} l(j)x_{t-j} + z_t, \quad (2.28)$$

$\{z_t\}$ is a stationary process uncorrelated with x_t and $\sum |l(j)| < \infty$.