LECTURE 7

Spectral Densities and Cumulants

We have already remarked on the theorem of Herglotz and what it implies for the representation of the covariance function of a stationary process X_k with finite second moments. Assume, for convenience, that $EX_k \equiv 0$. There is a result due to Cramér which gives a parallel representation of the process X_k itself as a Fourier–Stieltjes stochastic integral of a random process with orthogonal increments $z(\lambda)$,

$$(7.1) \qquad Ez(\lambda) \equiv 0, \qquad E\,dz(\lambda)\,\overline{dz(\mu)} = \delta(\lambda - \mu)\,dG(\lambda)$$

with δ the Kronecker delta

$$\delta(\lambda) = \begin{cases} 1 & \text{if } \lambda = 0, \\ 0 & \text{otherwise}. \end{cases}$$

If the process is real-valued,

$$dz(\lambda) = \overline{dz(-\lambda)},$$
$$dG(\lambda) = dG(-\lambda).$$

The representation of X_k is

$$X_k = \int_{-\pi}^{\pi} e^{ik\lambda}\,dz(\lambda).$$

Knowledge of the spectral distribution function or its derivative $g(\lambda)$ (assuming G absolutely continuous) is clearly of interest in a host of linear problems or Gaussian problems. However, in case of nonlinearity or of non-Gaussian character, higher order moments (assuming they exist) can convey additional information. Let

$$\varphi(t_1, \ldots, t_k) = E\exp\left\{ i\sum_{j=1}^{k} t_j X_j \right\} = \varphi(t).$$

The mixed moments

$$EX^v = EX_1^{v_1}, \dots, X_k^{v_k} = m_v$$

with the v_i nonnegative integers

$$v = (v_1, \dots, v_k), \qquad |v| = \sum_{j=1}^{k} v_j, \qquad v! = \prod_{j=1}^{k} v_j!,$$

if they exist up to an order n ($|v| \leq n$), can be identified as coefficients in the Taylor expansion of φ about zero,

$$\varphi(t) = \sum_{|v| \leq n} (it)^v m_v / v! + o(|t|^k).$$

Joint cumulants

$$c_v = \operatorname{cum}(X_1^{v_1}, \dots, X_k^{v_k})$$

are the corresponding coefficients in the expansion of log φ about 0

$$\log \varphi(t) = \sum_{|v| \leq n} (it)^v c_v / v! + o(|t|^k).$$

Existence of all moments up to order n is equivalent to existence of all moments up to order n. Cumulants up to order n can be expressed in terms of moments up to order n and the converse is also true. It is often much more convenient to deal with cumulants rather than moments. Notice that for jointly Gaussian variables all cumulants of order higher than the second are zero. One can show that if $(\nu_1, \dots, \nu_p)$ is a partition of the set of integers $\{1, 2, \dots, k\}$

$$E(X_1 \cdots X_k) = \sum_{\nu} C_{\nu_1} \cdots C_{\nu_p},$$

where C_v is the joint cumulant of the X's with subscripts in v. If μ_v is the mean of the product of the X's with subscripts in v, the inverse relation is given by

$$\operatorname{cum}(X_1, \dots, X_k) = \sum (-1)^{p-1} (p-1)! \mu_{\nu_1} \cdots \mu_{\nu_p}.$$

Assume that moments of order $k > 2$ exist. Then

$$m_k(j_1, \dots, j_k) = E[X_{j_1} \cdots X_{j_k}]$$

in the case of the strictly stationary sequence X_s will only depend on the time differences $j_2 - j_1, \dots, j_k - j_1$,

$$m_k(j_1, \dots, j_k) = r_k(j_2 - j_1, \dots, j_k - j_1).$$

A representation for r_k analogous to the result of Herglotz would be

$$r_k(\tau_1, \dots, \tau_{k-1}) = \int \overset{\pi}{\underset{-\pi}{\cdots}} \int \exp\left\{ i \sum_{\alpha=1}^{k-1} \tau_\alpha \lambda_\alpha \right\} dH_k(\lambda_1, \dots, \lambda_{k-1})$$

with H_k of bounded variation. However, this will generally not hold without strong enough assumptions such as, for example, cumulant mixing conditions of the type mentioned earlier. Even in such a case, the mass of H_k will be in part located in a singular manner on certain submanifolds if $k > 3$. We say $(\lambda_1, \dots, \lambda_k)$ lies on a proper submanifold if not only is

$$\sum_{j=1}^{k} \lambda_j \equiv 0 \mod 2\pi,$$

but also for a proper subset J of the set of integers $1, 2, \dots, k$,

$$\sum_{j \in J} \lambda_j \equiv 0 \mod 2\pi.$$

This will not be the case in such circumstances for the corresponding Fourier representation of the cumulants

$$\begin{aligned} c_k(\tau_1, \dots, \tau_{k-1}) &= \operatorname{cum}(X_0, X_{\tau_1}, \dots, X_{\tau_{k-1}}) \\ &= \int \overset{\pi}{\underset{-\pi}{\cdots}} \int \exp\left\{ \sum_{\alpha=1}^{k-1} \tau_\alpha \lambda_\alpha \right\} dG_k(\lambda_1, \dots, \lambda_{k-1}) \end{aligned}$$

with G_k of bounded variation. Actually with a cumulant mixing condition as strong as (5.2), G_k will be absolutely continuous with a density g_k. Let us illustrate these ideas in terms of a simple but interesting collection of linear models. Let ξ_k, $k = \dots, -1, 0, 1, \dots$, be independent identically distributed random variables with $E\xi_k \equiv 0$ and $E|\xi_k|^p < \infty$, $p > 2$ (p an integer). Let the corresponding cumulants of ξ_k be μ_s, $s = 0, 1, \dots, p$. It is clear that $\mu_0 = \mu_1 = 0$, $\mu_2 = \sigma^2(\xi)$. Let α_j be real weights with $\sum \alpha_j^2 < \infty$. Then

$$X_k = \sum_j \alpha_j \xi_{k-j}$$

is well defined. The spectral density of the process X_k is

$$g(\lambda) = \frac{\sigma^2}{2\pi} \left| \sum \alpha_j e^{ij\lambda} \right|^2 = \frac{1}{2\pi} \sum r_k e^{ik\lambda}, \tag{7.2}$$

with

$$r_k = \sum_k \alpha_k \alpha_{j+k} \sigma^2.$$

Let

$$\alpha(e^{-i\lambda}) = \sum \alpha_j e^{ij\lambda}.$$

One can then rewrite (7.2) as

$$g(\lambda) = \frac{\sigma^2}{2\pi} \left| \alpha(e^{-i\lambda}) \right|^2. \tag{7.3}$$

A relation analogous to (7.3) for cumulant spectra is

$$\begin{aligned}(7.4)\quad &\operatorname{cum}(dz(\lambda_1),\dots,dz(\lambda_k))\\ &\quad = \eta(\lambda_1+\cdots+\lambda_k)g_k(\lambda_1,\dots,\lambda_{k-1})\,d\lambda_1,\dots,d\lambda_{k-1},\end{aligned}$$

where

$$\eta(\lambda) = \begin{cases} 1 & \text{if } \lambda \equiv 0 \bmod 2\pi, \\ 0 & \text{otherwise.}\end{cases}$$

The sth order cumulant

$$\operatorname{cum}(X_t, X_{t+j_1},\dots,X_{t+j_{s-1}}) = \sum_k \alpha_k \alpha_{k+j_1} \cdots \alpha_{k+j_{s-1}} \mu_s$$

and because of this, one can see that the cumulant spectral density of order s is

$$\begin{aligned} &g_s(\lambda_1,\dots,\lambda_{s-1})\\ &\quad = (2\pi)^{-s+1} \sum_{j_1,\dots,j_{s-1}} \operatorname{cum}(X_t, X_{t+j_1},\dots,X_{t+j_{s-1}}) \exp\left(\sum_{k=1}^{s-1} i j_k \lambda_k\right)\\ &\quad = \mu_s (2\pi)^{-s+1} \alpha(e^{-i\lambda_1}) \cdots \alpha(e^{-i\lambda_{s-1}}) \alpha(\exp(i(\lambda_1+\cdots+\lambda_{s-1}))).\end{aligned}$$

It is clear from (7.3) that knowledge of the spectral density will only allow one to resolve $|\alpha(e^{-i\lambda})|$ but not the phase of $\alpha(e^{-i\lambda})$. In the case of a Gaussian process, because all higher (than second) order cumulants are zero, this means that $\alpha(e^{-i\lambda})$ is not identifiable. The way of resolving the question in the Gaussian case has been to make what is called a minimum phase assumption. For this to be meaningful, we have to assume that

$$\log g(\lambda) \in L.$$

Formally, let us just make a Fourier expansion of $\log g$,

$$\log g(\lambda) \sim \sum_{j=-\infty}^{\infty} b_j e^{-ij\lambda},$$

$$b_j = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log g(\lambda) e^{ij\lambda}\, d\lambda.$$

Then if we set

$$b(e^{-i\lambda}) = \tfrac{1}{2} b_0 + \sum_{j=1}^{\infty} b_j e^{-ij\lambda}$$

and

$$a(e^{-i\lambda}) = \exp\{b(e^{-i\lambda})\}$$

it is clear that

$$g(\lambda) = \left|a(e^{-i\lambda})\right|^2.$$

The minimum phase assumption is the assumption that

$$\frac{1}{\sqrt{2\pi}}\sigma\alpha(e^{-i\lambda}) = a(e^{-i\lambda})$$

and one just tries to estimate $a(e^{-i\lambda})$ (or its coefficients). Of course, in the Gaussian case one cannot distinguish between $a(e^{-i\lambda})$ and any other square root of $g(\lambda)$ so its just as well to estimate $a(e^{-i\lambda})$. In non-Gaussian cases, as we shall see, one can in principle resolve most of the phase information of $\alpha(e^{-i\lambda})$ under appropriate conditions. The minimum phase square root is tied up with the linear prediction problem for X_k and so has a natural aspect from that perspective.

It is worthwhile looking at the question just discussed in the simpler context of what is called an autoregressive moving average or ARMA scheme. Consider a linear system of equations

$$\sum_{j=1}^{p} \beta_j X_{t-j} = \sum_{k=0}^{q} \gamma_k \xi_{t-k} \tag{7.5}$$

with β_0, $\alpha_0 \neq 0$. An initial question is that of necessary and sufficient conditions for the existence of a stationary solution X_t. One can show that such a condition amounts to the polynomial

$$\beta(z) = \sum_{j=0}^{p} \beta_j z^j$$

having no zeros of absolute value one. If there is such a solution it is unique and given by

$$X_t = \int_{-\pi}^{\pi} e^{it\lambda}\frac{\gamma(e^{-i\lambda})}{\beta(e^{-i\lambda})}\, dz_{\xi}(\lambda), \tag{7.6}$$

where

$$\gamma(z) = \sum_{k=0}^{q} \gamma_k z^k$$

and $z_{\xi}(\lambda)$ is the random spectral function of the sequence $\{\xi_t\}$. In the discussion above it is assumed that $\beta(z)$, $\gamma(z)$ have no roots in common. In the context of ARMA processes, one can show that the minimum phase condition amounts to the assumption that $\beta(z)$, $\gamma(z)$ have all their zeros outside the unit disc in the complex plane. For convenience, assume that the zeros of $\beta(z)$ and $\gamma(z)$ are simple. The representation we obtain is still valid even if there are multiple zeros but the notation in the derivation would be somewhat more elaborate. Notice that if $|z_j| > 1$,

$$\begin{aligned}\left(e^{-i\lambda} - z_j\right)^{-1} &= (-z_j)^{-1}\left(1 - z_j^{-1}e^{-i\lambda}\right)^{-1}\\ &= (-z_j)^{-1}\sum_{k=0}^{\infty} z_j^{-k}e^{-ik\lambda}\end{aligned} \tag{7.7}$$

while if $|z_j| < 1$,

$$(7.8)\qquad \left(e^{-i\lambda} - z_j\right)^{-1} = e^{i\lambda}\left(1 - e^{i\lambda} z_j\right)^{-1} = \sum_{k=0}^{\infty} z_j^k e^{i(k+1)\lambda}.$$

If there is a stationary solution X_t of (7.5), it can be written

$$X_t = \sum_{k=-\infty}^{\infty} b_k \xi_{t-k},$$

where the coefficients b_k decrease in modulus exponentially fast as $|k| \to \infty$. This can be seen by making use of (7.6), (7.7) and (7.8). If all the roots of $\beta(z)$ have modulus greater than one, the representation of X_t in terms of the ξ_t's becomes one-sided,

$$X_t = \sum_{k=0}^{\infty} b_k \xi_{t-k}.$$

Moreover if $\gamma(z)$ also has all its roots with modulus greater than one, since

$$\xi_t = \int_{-\pi}^{\pi} e^{-it\lambda} \frac{\beta(e^{-i\lambda})}{\gamma(e^{-i\lambda})}\, dz_X(\lambda),$$

we have

$$\xi_t = \sum_{k-0}^{\infty} h_k X_{t-k},$$

with the weights h_k decreasing in absolute value exponentially fast. Thus, if $\beta(z)$, $\gamma(z)$ both have roots of absolute value greater than one, whether the ξ's are Gaussian or not, $L^2(\xi_t, t \le m)$ and $L^2(X_t, t \le m)$ (the spaces of random variables with finite second moment measurable with respect to the σ-fields generated by ξ_t, $t \le m$, and X_t, $t \le m$, respectively) are the same. This means that under these circumstances (the minimum phase condition)

$$X_t^* = \beta_0^{-1}\left\{ \sum_{k-1}^{q} \gamma_k \xi_{t-k} - \sum_{j=1}^{p} \beta_j X_{t-j} \right\}$$

is the best linear predictor of X_t in terms of the past X_τ, $\tau < t$, since the error

$$X_t - X_t^* = \beta_0^{-1} \gamma_0 \xi_t$$

is orthogonal to X_τ, $\tau < t$. However, it is not simply the *best linear predictor* of X_t in terms of the past. It is the *best predictor* of X_t in terms of the past in the sense of minimizing mean square error of prediction since $X_t - X_t^*$ is independent of $L^2(X_\tau,\ \tau \le t - 1)$. If the minimum phase condition is not satisfied, the best predictor may no longer be linear. A simple example illustrating this is given by the autoregressive system

$$X_t - 2X_{t-1} = -\eta_t$$

with the η_t's independent and

$$\eta_t = \begin{cases} 1 & \text{with probability } \frac{1}{2}, \\ 0 & \text{with probability } \frac{1}{2}. \end{cases}$$

Then

$$X_t = \sum_{k=1}^{\infty} \frac{1}{2^k} \eta_{t+k}.$$

The best predictor of X_t in terms of the past is

$$X_t^* = 2X_{t-1} \mod 1$$

and the prediction error is zero.

LEMMA. *Let $\{X_t\}$ be a non-Gaussian linear process. Assume that the generating independent random variables ξ_t have finite moments up to order $k(>2)$ with the kth cumulant $\mu_k \neq 0$. Further let*

$$\sum |j|\,|\alpha_j| < \infty$$

with $\alpha(e^{-i\lambda}) \neq 0$ for all λ. The function $\alpha(e^{-i\lambda})$ can then be identified in terms of observations on (X_t) alone up to the undetermined integer a in a factor $\exp(ia\lambda)$ *and sign of $\alpha(1) = \sum \alpha_j$.*

The cumulant spectral density of order k of the process X_t is $g_k(\lambda_1, \dots, \lambda_{k-1})$ [see (7.4)]. Now

$$\left\{ \frac{\alpha(1)}{|\alpha(1)|} \right\}^k \mu_k = (2\pi)^{k/2-1} \frac{g_k(0, \dots, 0)}{\{g(0)\}^{k/2}}.$$

Set

$$h(\lambda) = \arg\left\{ \alpha(e^{-i\lambda}) \frac{\alpha(1)}{|\alpha(1)|} \right\}.$$

Then

$$h(\lambda_1) + \cdots + h(\lambda_{k-1}) - h(\lambda_1 + \cdots + \lambda_{k-1})$$
$$= \arg\left[\left\{ \frac{\alpha(1)}{|\alpha(1)|} \right\}^k \mu_k^{-1} g_k(\lambda_1, \dots, \lambda_{k-1}) \right],$$

since

$$h(-\lambda) = -h(\lambda).$$

Also

$$h'(0) - h'(\lambda) = \lim_{\Delta \to 0} \{h(\lambda) + (k-2)h(\Delta) - h(\lambda + (k-2)\Delta)\}\{(k-2)\Delta\}^{-1}.$$

Notice that

$$h(\lambda) = \int_0^\lambda \{h'(u) - h'(0)\}\, du + c\lambda = h_1(\lambda) + c\lambda,$$

where $c = h'(0)$. Knowledge of the kth order cumulant spectral density g_k implies knowledge of $h_1(\lambda)$. The α_j's are real and so $h(\pi) = a\pi$ with a an integer. If $h_1(\pi)/\pi = \delta$, then $c = a - \delta$. a cannot be determined without additional assumptions. A change in a corresponds to reindexing the ξ_t's. The sign of $\alpha(1)$ cannot be resolved since multiplying the α_j's and ξ_j's by -1 does not change the process X_t.

Of course, under an assumption like that of an ARMA process, a is specified. Notice that the result above suggests the relevance of the estimation of both the spectral density g and the kth order cumulant spectral density g_k in the estimation of the function $\alpha(e^{-i\lambda})$. The estimation of g is useful in resolving the modulus of $\alpha(e^{-i\lambda})$ and that of g_k in determining the phase of $\alpha(e^{-i\lambda})$ up to the undetermined integer a and sign spoken of above.

There is a clear motivation for a result describing the asymptotic distribution of a class of spectral density estimates and we shall give one making use of the central limit theorem for strongly mixing triangular sequences described earlier. Assume that the process X_k has an absolutely summable covariance sequence

$$\sum_k |r_k| < \infty.$$

Covariance estimates

$$r_k^{(n)} = \frac{1}{n} \sum_{j=1}^{n-k} X_j X_{j+k}, \qquad k \geq 0,$$

with $r_{-k}^{(n)} = r_k^{(n)}$. The spectral density

$$f(\lambda) = \frac{1}{2\pi} \sum_k r_k e^{-ik\lambda}$$

and we shall consider estimates of the form

$$f_n(\lambda) = \frac{1}{2\pi} \sum_{k=-n+1}^{n-1} r_k^{(n)} \omega_k^{(n)} \cos k\lambda$$

with weights

$$\omega_k^{(n)} = a(kb_n), \qquad a(0) = 1.$$

The function $a(x)$ is assumed to be continuous at zero, bounded and symmetric,

$$a(x) = a(-x).$$

Under these circumstances if $b_n \to 0$, one can show that the estimates are asymptotically unbiased.

LEMMA. *Suppose*

$$h(u) = \frac{1}{2\pi} \sum h_k e^{-iku}, \qquad \sum |h_k| < \infty.$$

Also let $a(x)$ *be piecewise continuous, continuous at* 0 *with* $a(0) = 1$, *symmetric as well as* $a(x) = O(|x|^{-(1/2)-\varepsilon})$ *for some* $\varepsilon > 0$ *as* $|x| \to \infty$. *Given*

$$W_n(u) = \frac{1}{2\pi} \sum \omega_k^{(n)} e^{-iku}$$

with $\omega_k^{(n)} = a(kb_n)$, *we have*

$$b_n \int_{-\pi}^{\pi} W_n^2(u + \lambda) g(u)\, du \to g(\lambda) \int W^2(u)\, du$$

as $b_n \to 0$, *where*

$$W(u) = \frac{1}{2\pi} \int a(u) e^{-iu\alpha}\, d\alpha.$$

Also

$$\int_{-\pi}^{\pi} W_n(u + \lambda) W_n(u + \mu) g(u)\, du = o(b_n^{-1})$$

as $b_n \to 0$, *if* $\lambda \neq \mu$, $|\lambda - \mu| < 2\pi$.

The technical lemma above is useful in determining asymptotic behavior of the covariance properties of spectral estimates at different frequencies. The argument for the lemma can be carried through first for step functions $a(\cdot)$ with finite support and then using an approximation argument.

We wish to obtain the following result.

THEOREM. *Let* $X = \{X_n\}$ *be a strictly stationary strongly mixing process with* $EX_j = 0$. *Let the cumulant functions of* X *up to order* 8 *be absolutely summable. Let the weights* $\omega_k^{(n)}$ *be given by a function* $a(\cdot)$ *that is piecewise continuous, continuous at* 0 *with* $a(0) = 1$, $a(u) = a(-u)$ *and such that* $xa(x)$ *is bounded. Then* $[f_n(\lambda) - Ef_n(\lambda)](nb_n)^{1/2}$ *is asymptotically normally distributed with mean zero and variance*

$$\frac{2\pi(1 + \eta(\lambda))}{nb_n} f^2(\lambda) \int W^2(\alpha)\, d\alpha,$$

where

$$\eta(\lambda) = \begin{cases} 1 & \text{if } \lambda = k\pi,\ k \text{ integer}, \\ 0 & \text{otherwise}. \end{cases}$$

If we consider jointly spectral estimates at distinct frequencies in $[0, \pi]$, *the centered and normalized estimates are asymptotically independent and normal with mean zero and variances given above with* $f(\lambda)$ *computed at the appropriate frequencies* λ.

First consider $a(\cdot)$ with finite support. The support is taken to be $[-1, 1]$ for convenience but the argument is the same for any interval. We wish to first show one can replace $f_n(\lambda)$ by

$$\tilde{f}_n(\lambda) = \frac{1}{2\pi} \sum_{k=-c(n)}^{c(n)} \frac{1}{n} \sum_{j=1}^{n} X_j X_{j+k} \omega_k^{(n)} \cos k\lambda$$

with $c(n) = b(n)^{-1}$. Now

$$2\pi n \tilde{f}_n(\lambda) = \sum_{u=1}^{n} Y_u^{(n)}$$

with

$$Y_u^{(n)} = \sum_{k=-c(n)}^{c(n)} X_u X_{u+k} \omega_k^{(n)} \operatorname{cov} k\lambda .$$

But

$$\sigma^2\big(f_n(\lambda) - \tilde{f}_n(\lambda)\big) \le \frac{1}{\pi^2 n^2} \sigma^2 \left[\sum_{k-1}^{c(n)} \sum_{j-n-k}^{n} X_j X_{j+k} \omega_k^{(n)} \cos k\lambda \right]$$

with the right side of the previous inequality less than

$$\frac{1}{\pi^2 n^2} \sum_{k,k'-1}^{c(n)} \sum_{j-n-k}^{n} \sum_{j'=n-k'}^{n} \Big\{ |r_{j-j'} r_{j-j'+k-k'}| + |r_{j-j'-k'} r_{j-j'+k}| + \big| r^{(4)}_{k,j'-j,j'+k'-j} \big| \Big\} \big| \omega_k^{(n)} \omega_{k'}^{(n)} \big|$$

$$(7.9) \qquad \le \frac{1}{\pi^2 n^2} \sum_{k,k'=1}^{c(n)} \sum_{s} \min(k', k) \Big\{ |r_s|\,|r_{s+k'-k}| + |r_{s-k'} r_{s+k}| + \big| r^{(4)}_{k,s,s+k'} \big| \Big\} \big| \omega_k^{(n)} \big| \, \big| \omega_{k'}^{(n)} \big| ,$$

where $r^{(4)}$ denotes the fourth order cumulant function. Conditions assumed relative to the weight $\omega_k^{(n)}$ and the function $a(\cdot)$ imply bounded weights and

$$|k| \big| \omega_k^{(n)} \big| \le L b_n^{-1}$$

for all k and some constant L. Absolute summability of cumulants of second and fourth order together with (7.9) implies that

$$(7.10) \qquad \sigma^2\big(f_n(\lambda) - \tilde{f}_n(\lambda)\big) = O\big((nb_n)^{-2}\big) = o\big((nb)^{-1}\big).$$

This is of smaller order of magnitude than $\sigma^2(\tilde{f}(\lambda))$. Now

$$\sigma^2\left(\sum_{u=1}^{m} Y_u^{(n)}\right) = h_n(m(n))$$

$$= \sum_{u,u'=1}^{m} \sum_{k,k'=-c(n)}^{c(n)} \{r_{u-u'}r_{u+k-u'-k'} + r_{u'-u+k'}r_{u'-u-k}$$

$$+ r^{(4)}_{k,u'-u,u'-u+k'}\}\cos k\lambda \cos k'\lambda\, \omega_k^{(n)}\omega_{k'}^{(n)}$$

$$= (1) + (2) + (3),$$

where $m = m(n) \le n$ with $b_n^{-1} = o(m(n))$ as $n \to \infty$. First

$$(1) = \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \frac{\sin^2(m/2)(\alpha-\beta)}{\sin^2(1/2)(\alpha-\beta)} f(\alpha) f(\beta) \left| \sum_{v=-c(n)}^{c(n)} \omega_v^{(n)} \cos v\lambda e^{iv\beta}\right|^2 d\alpha\, d\beta.$$

Also

$$\max_{|u|\le Am^{-1}} \left| \frac{\int_{-\pi}^{\pi} W_n(v)W_n(v+u)\,dv}{\int_{-\pi}^{\pi} W_n^2(u)\,du} - 1\right| \to 0$$

as $m = m(n) \to \infty$ for each $A > 0$ because $c(n) = b_n^{-1} = o(m(n))$. One can show that

$$(1) = 2\pi m(1+o(1))\int_{-\pi}^{\pi} f^2(\beta)\left|\tfrac{1}{2}W_n(\beta-\lambda) + \tfrac{1}{2}W_n(\beta+\lambda)\right|^2 d\beta.$$

The lemma implies

$$(7.11) \qquad (1) = \pi m(1+\eta(\lambda)+o(1))\, f^2(\lambda)\int_{-\pi}^{\pi} W_n^2(v)\,dv.$$

One can show that term (2) asymptotically has the same behavior. Since

$$|(3)| \le m \sum_{k,s,k'} \left|r^{(4)}_{k,s,k'}\right|,$$

it is of smaller order than (7.11). It is clear that

$$\int_{-\pi}^{\pi} W_n^2(u)\,du = (1+o(1))b_n^{-1}\int W^2(u)\,du.$$

Notice that $\sigma^2(f_n(\lambda))$ and $\sigma^2(\tilde{f}_n(\lambda))$ have the same asymptotic behavior as $n \to \infty$ by (7.10). Because of the summability of second and fourth order cumulants, if $m(n) = o(n)$ and $k(n)m(n) = n$, then

$$k(n)h_n(m(n)) \simeq h_n(n).$$

In order to apply the central limit theorem we have to show that

$$\sigma^{-4}\left(\sum_{u=1}^{m} Y_u^{(n)}\right) E\left|\sum_{u=1}^{m}\left(Y_u^{(n)} - EY_v^{(n)}\right)\right|^4 = O(1).$$

But

$$E\left|\sum_{u=1}^{m}\left(Y_u^{(n)} - EY_u^{(n)}\right)\right|^4 = \sigma^4\left(\sum_{u=1}^{m} Y_u^{(n)}\right) + \underset{4}{\mathrm{cum}}\left(\sum Y_u^{(n)}\right).$$

We use the multilinear character of the cumulant function and observe that

$$\begin{aligned} &\mathrm{cum}\, 4\left(Y_{u_1}^{(n)}, Y_{u_2}^{(n)}, Y_{u_3}^{(n)}, Y_{u_4}^{(n)}\right) \\ &\quad = \left(\prod_{i=1}^{4} \omega_{k_i}^{(n)} \cos k_i\lambda\right) \sum_{v} \mathrm{cum}(X_s, s \in v_1) \cdots (\mathrm{cum}\, X_s, s \in v_p), \end{aligned} \tag{7.12}$$

where the sum is over all indecomposable partitions of the table

$$\begin{matrix} x_{u_1} x_{u_1+k_1} \\ x_{u_2} x_{u_2+k_2} \\ x_{u_3} x_{u_3+k_3} \\ x_{u_4} x_{u_4+k_4}. \end{matrix}$$

We just analyze one of the many indecomposable partitions consisting entirely of pairs. It leads to the sum

$$\sum_{k_i}\sum_{u_i} r_{u_1-u_2} r_{u_3-u_4} r_{u_3-u_1+k_3-k_1} r_{u_4-u_2+k_4-k_2} \prod_{i=1}^{4} \omega_{k_i}^{(n)} \cos k_i\lambda.$$

Let $a = u_1 - u_2$, $b = u_3 - u_4$, $\alpha = k_3 - k_1$, $\beta = k_4 - k_2$. The expression is bounded by

$$\sum_{a, b, u_1, u_3, k_1, k_2, \alpha, \beta} |r_a|\,|r_b|\,|r_{u_3-u_1+\alpha}|\,|r_{u_3-u_1+b+\alpha+\beta}|\,|\omega_{k_1}^{(n)}|\,|\omega_{k_1+\alpha}^{(n)}|\,|\omega_{k_2}^{(n)}|\,|\omega_{k_2+\beta}^{(n)}|.$$

Sum over k_1, k_2 first to get the bound

$$b_n^{-2} \sum_{a, b, u_1, u_3, \alpha, \beta} |r_a|\,|r_b|\,|r_{u_3-u_1+\alpha}|\,|r_{u_3-u_1-b+\alpha+\beta}|.$$

Now sum over α, β to get

$$b_n^{-2} \sum_{a, b, u_1, u_3} |r_a|\,|r_b|.$$

The sum over a, b and then u_1, u_3 yields the bound

$$b_n^{-2} m^2.$$

This is of the order of magnitude of

$$\sigma^{-4}\left(\sum_{u=1}^{m} Y_u^{(n)}\right).$$

The sums over all other terms stemming from (7.12) gives us something of this order of magnitude or less.

There is a large literature concerned with parameter estimation for Gaussian sequences [see Brockwell and Davis (1987)]. For convenience we consider the case of a stationary Gaussian autoregressive scheme ($EX_t = E\xi_t \equiv 0$)

$$X_t = \sum_{i=1}^{p} a_i X_{t-i} + \xi_t \tag{7.13}$$

with independent identically distributed residuals ξ_t. Typically remarks can be made for moving averages or ARMA schemes that are similar to those that follow. To avoid nonidentifiability in estimating the coefficients a_i, one assumes that the polynomial

$$a(z) = 1 - \sum_{j=1}^{p} a_j z^j$$

has all its zeros with modulus greater than one (the minimum phase condition). This assumption need not be made if the process X_t is non-Gaussian since the nonidentifiability does not arise then. If one minimizes

$$\sum_{t=p+1}^{n} \left(X_t - a_1 X_{t-1} - \cdots - a_p X_{t-p}\right)^2,$$

estimates of $a_1, \dots, a_p, \sigma^2$ $(\alpha_1, \dots, \alpha_p, s^2)$ are obtained which satisfy

$$\begin{aligned} &\sum_{t=p+1}^{n} X_t X_{t-j} - \sum_{i=1}^{p} \alpha_i \sum_{t=p+1}^{n} X_{t-i} X_{t-j} = 0, \qquad j = 1, \dots, p, \\ &s^2 = \frac{1}{n} \sum_{t=p+1}^{n} \left(X_t - \alpha_1 X_{t-1} - \cdots - \alpha_p X_{t-p}\right)^2 \end{aligned} \tag{7.14}$$

and are $n^{-1/2}$ consistent. Let σ^2 be the variance of the ξ_t's. The random variables $n^{1/2}(\alpha_i - a_i)$, $i = 1, \dots, p$, are asymptotically normal with mean zero and covariance

$$\sigma^2 R^{-1}.$$

The $p \times p$ matrix R is the covariance matrix of $X_1, \dots, X_p$. Also $n^{1/2}(s^2 - \sigma^2)$ is asymptotically uncorrelated with the α_i's and is asymptotically normal with mean zero and variance $\mu_4 = E\xi_t^4$.

If the process X_t of (7.13) is non-Gaussian with the ξ_t process orthonormal, strongly mixing in an appropriate sense and satisfying suitable moment conditions, one could still consider the same estimates of the coefficients a_i in the minimum phase context. The same asymptotic results would hold for the estimates α_i of a_i [see Rosenblatt (1985)]. However, if one had more information about the distribution, for example, of the ξ_t's, one would expect to be able to get even better estimates.

Some interesting results of Kreiss (1987) that indicate how one can obtain asymptotically optimal estimates of the a_i's via an adaptive procedure making use of probability density estimates will be mentioned. Assume that the autoregressive scheme is still minimum phase and that the ξ_t's are independent with a density $f > 0$ that is absolutely continuous and with finite Fisher information

$$I(f) = \int (f'/f)^2 f\,dx < \infty.$$

First a procedure is considered assuming that the density f is known. Let

$$\Delta_n = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \dot{\varphi}(\xi_j) X(j-1),$$

where

$$\dot{\varphi} = -f'/f, \qquad \xi_j = X_j - a^T X(j-1)$$

and

$$a = (a_1, \dots, a_p), \qquad X(s) = (X_s, \dots, X_{s-p+1}).$$

The basic idea is to get an initial estimate of the a_i's that is $n^{1/2}$ consistent (and discrete). The estimate is then adjusted so as to get one that is asymptotically optimal. If such an estimate α of a is used to compute estimates of ξ_j and Δ_n, we shall refer to them as $\xi_j(\alpha)$ and $\Delta_n(\alpha)$. Such an initial estimate is provided by the solution α of the system (7.14). Kreiss then shows that the estimate

$$\hat{\alpha} = \alpha + n^{-1/2} \frac{R^{-1}}{I(f)} \Delta_n(\alpha)$$

is asymptotically optimal as $n \to \infty$. $n^{1/2}(\hat{\alpha} - a)$ is asymptotically normal as $n \to \infty$ with mean zero and covariance matrix $R^{-1}/I(f)$. To get an adaptive procedure, it is clear that one needs an estimate of the unknown density $f(x)$. Kreiss makes use of the initial $\sqrt{n}$ consistent estimate of the a_i's to estimate the residuals ξ_t and in terms of these gets an appropriate estimate of the unknown f that is adequate for his purposes. The upgrading of a $\sqrt{n}$ consistent estimate to one that is asymptotically optimal is a bit reminiscent qualitatively of the discussion in Lehmann [(1983), page 422] in another context.

The beginning of an analysis of a nonminimum phase non-Gaussian optimal estimate of coefficients for finite parameters schemes can be seen in Breidt, Davis, Lii and Rosenblatt (1990).

From the earlier discussion of the transfer function $\alpha(e^{-i\lambda})$ when observing a non-Gaussian linear process, it is clear that if the cumulant μ_s for some

$s > 2$ is nonzero, one would be able to estimate $\alpha(e^{-i\lambda})$ up to an indeterminate sign ± 1 and a phase factor $\exp(ia\lambda)$ (with a integral) by the lemma. This would require estimating the spectral density $g(\lambda)$ and the sth order cumulant spectral density $g_s(\lambda_1, \dots, \lambda_{s-1})$. Notice that if one only observes the linear process X_t and estimates $\alpha(e^{-i\lambda})$ (assumed nonzero), it is then possible to estimate the ξ_k process and so deconvolve the X_t process. Deconvolution problems like this arise in a geophysical context. The constants α_j are considered descriptive of a disturbance passing through a layered medium in a model for types of seismic exploration. The random values ξ_j are thought of as the reflectivity of slabs in the layered medium. At times observed data is definitely non-Gaussian and one wishes to deconvolve the data X_t, estimating the α_j's and ξ_j's. A discussion of some of these questions can be found in Donoho (1981) and Wiggins (1978). A detailed analysis of the deconvolution and estimation of the transfer function is given in Lii and Rosenblatt (1982). This analysis requires estimation of a third or fourth order cumulant spectral density for the sequence X_t. Estimates of such cumulant spectral densities using the fast Fourier transform are considered in Brillinger and Rosenblatt (1967). The results on asymptotic normality as remarked in Lecture 5 require existence of all moments and an infinite number of cumulant summability conditions. They also make use of values of smoothed versions of periodogram-like functions (computed from the finite Fourier transform) at values of the arguments of the form $2\pi s/N$ with s integral and N the sample size. The discussion of spectral estimates (second order) given earlier required existence of moments up to order 8 only with a finite number of corresponding cumulant summability conditions. Corresponding results for higher order cumulant spectral estimates can be found in Lii and Rosenblatt (1990).

We also note that higher order cumulant spectra have been used in the analysis of a number of nonlinear problems. Third order spectra have been used in the model of homogeneous turbulence [see Batchelor (1953) and Lii, Rosenblatt and Van Atta (1976)] as a gauge to measure the nonlinear transfer of energy between wave number vectors in a harmonic analysis of a turbulent velocity field.