Autoregressive Processes

Robert L. Wolpert
Department of Statistical Science
Duke University, Durham, NC, USA

Vsn 2.2, 2014/03/29

1 Moving Averages

The simplest stationary time series beyond the iid case is the *moving average*, denoted $X \sim \mathsf{MA}(q)$, of the form

$$X_t = \sum_{i=0}^{q} b_i \zeta_{t-i}, \qquad t \in \mathbb{Z} \tag{1a}$$

for some integer $q \ge 0$, vector $b \in \mathbb{R}^{q+1}$, and iid sequence $\{\zeta_t\}$. Without any loss of generality we may take $\beta_0 = 1$. If $\{\zeta_t\} \subset L_1$ with mean $\mu = \mathsf{E}\zeta_t$ then X_t is in L_1 too, with mean $\mathsf{E}X_t = \mu \sum b_i$. If $\{\zeta_t\} \subset L_2$ with variance $\sigma^2 = \mathsf{E}(\zeta_t - \mu)^2$ then X_t is in L_2 too, with isotropic covariance

$$c(s,t) = \mathsf{Cov}(X_s, X_t) = \sigma^2 \Big\{ \sum_{j=0}^{q-|t-s|} b_{|t-s|+j} b_j \Big\} \tag{1b}$$

for $|t-s| \le q$, zero for $|t-s| > q$, that depends only on $|t-s|$. If an infinite sequence $\{b_j : j \ge 0\} \subset \mathbb{R}$ is absolutely summable then (1a) still defines a stationary L_1 sequence for iid $\{\zeta_i\} \subset L_1$ and, for $\{\zeta_i\} \subset L_2$, (1a) defines an L_2 process with covariance given by the now infinite but absolutely convergent sum (1b) (note $\sum_j b_j^2 < \infty$ if $\{b_j\}$ is absolutely summable).

The first-order autocorrelation $\rho = c(0,1)/c(0,0)$ for an $\mathsf{MA}(1)$ process is

$$\rho = \frac{c(0,1)}{c(0,0)} = \frac{b_1}{1 + b_1^2}$$

This takes a maximum value of $1/2$ at $b_1 = 1$ and a minimum of $-1/2$ at $b_1 = -1$, so always $\rho \in [-\frac{1}{2}, \frac{1}{2}]$. This illustrates that the possible values of the correlation function for an $\mathsf{MA}(q)$ process are rather limited. For any value of $0 < \rho < \frac{1}{2}$ there are two different values of b_1 that lead to autocorrelation ρ, one each in the intervals $(0,1)$ and $(1,\infty)$ (solutions to a quadratic equation).

1.1 Continuous Time

A stationary stochastic process indexed by $t \in \mathbb{R}$ (not just $\mathbb{Z}$) may be constructed in a way similar to that of (1a). Let $\zeta(ds)$ be a random measure on the Borel sets $\mathcal{B}(\mathbb{R})$ that assigns independent random variables $\zeta(A_i)$ to bounded disjoint Borel sets $A_i \subset \mathbb{R}$ and that is "stationary" in the sense that the joint distribution of any finite collection $\{\zeta(A_i)\}$ is identical to that of $\{\zeta(A_i + h)\}$ for each $h \in \mathbb{R}$, where $A + h = \{x + h : x \in A\}$ denotes the set A translated by h. If $\zeta(A) \in L_1$ for each bounded Borel A, and if $b \in L_1(\mathbb{R})$, then

$$X_t = \int_{\mathbb{R}} b(t - s)\zeta(ds), \qquad t \in \mathbb{R} \tag{2a}$$

determines a stationary L_1 stochastic process. If $b \in L_2(\mathbb{R})$ and if $\mathsf{E}\zeta(A)^2 < \infty$ for bounded A, then also $X_t \in L_2$ with covariance $c(s,t) = c(s - t)$ for

$$c(h) = \sigma^2 \int_{\mathbb{R}} b(t - h)b(t)\,dt \tag{2b}$$

where $\sigma^2 = \mathsf{V}\big(\zeta([0,1])\big)$. The conditions imposed on $\zeta(ds)$ imply that each $\zeta(A)$ must have an infinitely-divisible distribution, such as the normal, gamma, negative binomial, or Poisson; we will see more of this in a few weeks.

2 Autoregressive

Now let's consider interchanging the roles of X_t and ζ_t in (1a), and consider a process X_t for which

$$\sum_{j=0}^{p} a_j X_{t-j} = \zeta_t \qquad t \in \mathbb{Z} \tag{3}$$

for some integer $p \geq 0$, vector $a \in \mathbb{R}^{p+1}$, and iid sequence $\{\zeta_t\}$. Again we may take $a_0 = 1$ with no loss of generality. To simplify things a bit we will take $\mathsf{E}\zeta_t = 0$ and suppose $\sigma^2 = \mathsf{E}\zeta_t^2 < \infty$. Such a process is called "autoregressive of order p", denoted $X \sim \mathsf{AR}(p)$.

2.1 $\mathsf{AR}(1)$ Processes

The most commonly occurring example is $\mathsf{AR}(1)$, for which (3) may be rewritten

$$X_t = \rho X_{t-1} + \zeta_t \tag{4}$$

where $\rho = -a_1$ is the one-step autocorrelation and, by recursion, the general autocorrelation is

$$\mathsf{Corr}(X_s, X_t) = \rho^{|t-s|}$$

Applying (4) repeatedly n times leads to

$$X_t = \rho^n X_{t-n} + \sum_{0 \leq j < n} \rho^j \zeta_{t-j}$$

or, in the limit as $n \to \infty$, to the $\mathsf{MA}(\infty)$ representation

$$= \sum_{j=0}^{\infty} \rho^j \zeta_{t-j}$$

provided that $|\rho| < 1$ so the sequence $b_i = \rho^i$ for $i \geq 0$ will be absolutely summable. The autocovariance $\mathsf{Cov}(X_s, X_t) = c(|t-s|)$ now follows from (1b):

$$c(h) = \sigma^2 \sum_{j=0}^{\infty} \rho^{h+2j} = \frac{\sigma^2}{1-\rho^2}\,\rho^h, \qquad h \in \mathbb{Z}_+ \tag{5}$$

For $0 < \rho < 1$ this decreases monotonically, exponentially fast, while for $-1 < \rho < 0$ it alternates sign and for $\rho = 0$ the sequence $X_t = \zeta_t$ is iid. Unlike $\mathsf{MA}(q)$ processes for $q \geq 1$, $X_t \sim \mathsf{AR}(1)$ is a Markov process: the conditional distribution of X_t for $t > s$, given $\mathcal{F}_s = \sigma\{X_u : u \leq s\}$, coincides with that of X_t given X_s. In the important special case of $\{\zeta_t\} \overset{\text{iid}}{\sim} \mathsf{No}(0, \sigma^2)$, for $s \leq t$ this is

$$X_t \mid \mathcal{F}_s \sim \mathsf{No}\big(\rho^{t-s} X_s,\ \sigma^2\big).$$

2.2 Characteristic Polynomials

Consider the left-shift or lag operator L given by $LX_t = X_{t-1}$, and its positive integer powers (under composition), $L^j X_t = X_{t-j}$. Formally we may rewrite (3) as

$$P(L)X_t = \zeta_t \tag{6}$$

in terms of L and the *characteristic polynomial*

$$P(z) = \sum_{j=0}^{p} a_j\, z^j. \tag{7}$$

Any pth-order polynomial like $P(z)$ which satisfies $P(0) = 1$ can be factored over the complex field as the product

$$= \prod_{n=1}^{p} (1 - z/r_n)$$

of p factors, where $\{r_n\} \subset \mathbb{C}$ are the p (not necessarily distinct) complex roots of the equation $P(r_n) = 0$. The multiplicative inverse of $P(z)$ for z unequal to any of these roots can thus be written

$$1/P(z) = \prod_{n=1}^{p} \frac{1}{1 - z/r_n} \tag{8}$$

which has a power series expression

$$= \sum_{i=0}^{\infty} b_i z^i$$

with a radius of convergence $|z| < R \equiv \min_{n \leq p} |r_n|$. Provided that radius R exceeds one, *i.e.*, provided that all the roots of $P(z)$ lie outside the unit disk in $\mathbb{C}$, the sequence b_i will be absolutely summable and the MA(∞) representation suggested by formally multiplying (6) by $1/P(L)$ from (8) is valid:

$$X_t = \sum_{i=0}^{\infty} b_i \, \zeta_{t-i}. \tag{9}$$

Using this we can make least-squares predictions or, in the Gaussian case, conditional expectations:

$$\mathsf{E}[X_{t+h} \mid \mathcal{F}_t] = \sum_{i=0}^{\infty} b_{i+h} \zeta_{t-i}$$

which will be shown in Section (2.3.3) to be a linear function of just the p most recent values of X, $\{X_s : \ t-p < s \le t\}$. Because of this feature, the process X_t is called "p-Markov".

If we take any one root— say, r_p— and consider the process $Y_t \equiv X_t - X_{t-1}/r_p = (1 - L/r_p)X_t$, we see that $Y_t \sim \mathsf{AR}(p-1)$ characteristic polynomial $\tilde{P}(z) = \prod_{1 \le n < p}(1 - z/r_n)$ of order $(p-1)$; many properties of $X_t \sim \mathsf{AR}(p)$ can be proved by induction using this fact.

2.3 Yule-Walker Equations

If we multiply Eqn (3) by X_{t-h} for $h \ge 0$ and take expectations, we get

$$\begin{aligned} \mathsf{E}\Big\{ \Big[\sum_{j=0}^{p} a_j X_{t-j} \Big] X_{t-h} \Big\} &= \mathsf{E}\Big\{ \ [\zeta_t] \ X_{t-h} \Big\} \\ \sum_{j=0}^{p} a_j c(h-j) &= \begin{cases} \sigma^2 & h = 0 \\ 0 & h \ge 1 \end{cases} \end{aligned} \tag{10}$$

since ζ_t is independent of X_{t-h} for $h > 0$. This result, the Yule-Walker equation, allows us to derive the covariance function $c(h)$ for $\mathsf{AR}(p)$ processes or, conversely, to infer $\{a_j\}$ from specified or estimated values of the autocovariance function $c(\cdot)$.

If $r_n \in \mathbb{R}$ is a real root of $P(z)$ of (7) then $c_n(x) := r_n^{-|x|}$ satisfies (10) for $h \ge p$, because

$$\sum_{j-0}^{p} a_j \, c_n(h-j) = \sum_{j-0}^{p} a_j \, r_n^{-|h-j|} = \sum_{j-0}^{p} a_j \, r_n^{j-h} = r_n^{-h} P(r_n) = 0.$$

If $P(z)$ has p distinct real roots then some unique linear combination

$$c(h) = \sum_{n=1}^{p} \alpha_n r_n^{-|h|}$$

of the p linearly-independent functions $\{c_n(\cdot)\}$ also satisfies the p linear conditions (10) for $0 \le h < p$ and so satisfies (10) for all integers $h \in \mathbb{Z}_+$, a generalization of (5) from $\mathsf{AR}(1)$ to the AR(p). What if the roots aren't distinct, or if some aren't real? Both possibilities can arise for $p = 2$; let's look at that case first.

2.3.1 AR(2) Processes

For the AR(2) process

$$X_t + a_1 X_{t-1} + a_2 X_{t-2} = \zeta_t$$

there are three possibilities for the characteristic polynomial

$$P(z) = 1 + a_1 z + a_2 z^2 = (1 - z/r_1)(1 - z/r_2),$$

namely

1)	$a_1^2 > 4a_2$	Unequal real roots	$r_1 \neq r_2,\ r_1, r_2 \in \mathbb{R}$
2)	$a_1^2 = 4a_2$	Equal real roots	$r_1 = r_2 \in \mathbb{R}$
3)	$a_1^2 < 4a_2$	Conjugate Pair roots	$r_1 = \bar{r}_2 \in \mathbb{C}\backslash\mathbb{R}$

with very different behavior for sample paths and the covariance function.

In case 1) with distinct real roots we have just seen that the autocovariance function will be of the form $c(h) = \alpha_1 r_1^{-|h|} + \alpha_2 r_2^{-|h|}$, a linear combination of two geometrically-decreasing functions, each either positive if each $r_n > 1$ or alternating if $r_n < -1$.

In case 2) with equal real roots $r_1 = r_2 = r$, the characteristic polynomial is $P(z) = (1 - z/r)^2$ and the Yule-Walker equations (10) become

$$c(h) - 2r^{-1}c(h-1) + r^{-2}c(h-2) = \sigma^2 \mathbf{1}_{\{h=0\}}$$

Both $c_1(h) := r^{-|h|}$ and $c_2(h) := |h| r^{-|h|}$ satisfy this equation, so a unique linear combination $c(h) = \alpha_1 r^{-|h|} + \alpha_2 |h| r^{-|h|}$ will satisfy (10) for all $h \in \mathbb{Z}$.

Finally, in case 3), the roots of $P(z) = 1 - z[2\cos(\theta)/r] + z^2/r^2$ will be a complex conjugate pair $r_1 = re^{i\theta}$, $r_2 = re^{-i\theta}$ for some $r > 1$ and $0 < \theta < \pi$, so (10) becomes

$$c(h) - 2\cos(\theta) r^{-1} c(h-1) + r^{-2} c(h-2) = \sigma^2 \mathbf{1}_{\{h=0\}}.$$

This is satisfied for $h \geq 2$ by the linearly independent pair $c_1(h) := r^{-|h|}\cos(h\theta)$ and $c_2(h) := r^{-|h|}\sin(|h|\theta)$, so a unique linear combination $c(h) = \alpha_1 r^{-|h|}\cos(h\theta) + \alpha_2 r^{-|h|}\sin(|h|\theta)$ will satisfy (10) for all $h \in \mathbb{Z}$.

2.3.2 AR(p)

In general, the degree p polynomial $P(z)$ of (7) with real coefficients will have p roots $\{r_n\} \subset \mathbb{C}$, counted according to multiplicity, each either real or part of a complex-conjugate pair. The autocovariance function $c(h)$ will be a linear combination of p terms, each of the form $r_n^{-|h|}|h|^k$ for some $k \geq 0$ with $k = 0$ for isolated real roots, $k > 0$ for repeated ones, and $|r_n|^{-|h|}|h|^k \cos(h\theta_n)$, $|r_n|^{-|h|}|h|^k \sin(|h|\theta_n)$ for (repeated, if $k > 0$) complex conjugate pairs.

2.3.3 Prediction

By induction an AR(1) process satisfying $X_t + a_1 X_{t-1} = \zeta_t$ can be written

$$X_{t+h} = \rho^h X_t + \sum_{n=0}^{h-1} \rho^n \zeta_{t+h-n}$$

for $h \in \mathbb{Z}_+$, where $\rho := -a_1$, so the h-step-ahead prediction is

$$\mathsf{E}[X_{t+h} \mid \mathcal{F}_t] = \rho^h X_t.$$

For an $\mathsf{AR}(\mathrm{p})$ process with any $p \geq 0$ the *one*-step-ahead prediction is just as simple,

$$X_{t+1} = -\sum_{j=1}^{p} a_j X_{t+1-j} + \zeta_{t+1},$$
$$\mathsf{E}[X_{t+1} \mid \mathcal{F}_t] = -\sum_{j=1}^{p} a_j X_{t+1-j}. \tag{11}$$

Looking ahead $h > 1$ steps is a little more work. One alternative is to apply (11) recursively: first to find $\hat{X}_{t+1} = \mathsf{E}[X_{t+1} \mid \mathcal{F}_t]$; then with $\hat{X}_{t+1}$ playing the role of X_{t+1} to find $\hat{X}_{t+2} = \mathsf{E}[X_{t+2} \mid \mathcal{F}_t]$, and so on. Another is to recognize that $\mathsf{E}[X_{t+h} \mid \mathcal{F}_t]$ will be a linear combination of $\{c(h-j) : 0 \leq j < p\}$ for all $h \geq 0$, with coefficients that depend only on $\{X_{t-j} : 0 \leq j < p\}$; from Sections (2.3.1) & (2.3.2) we see that this will be a linear combination of terms of the form $r_n^{-|h|}$ for distinct real roots, $r_n^{-|h|}|h|^k$ for repeated real roots, and terms like $r^{-|h|}\cos(h\theta)$, $r^{-|h|}\sin(|h|\theta)$ for complex roots. A little linear algebra will identify the coefficients in this p-dimensional linear system.

The class of $\mathsf{AR}(\mathrm{p})$ processes were introduced independently by Yule (1927) and Walker (1925), in both cases with $p = 2$ and motivated by a desire to model oscillatory natural phenomena (sunspots and the SMO, respectively) in "case 3)" above with complex roots.

3 Spectra

Let $\gamma(\omega)$ be a nonnegative integrable function on $[0, \pi]$, and let $\mathcal{Z}_1(d\omega)$ and $\mathcal{Z}_2(d\omega)$ be independent mean-zero Gaussian random measures on $[0, \pi]$ with

$$\mathsf{E}\left\{\mathcal{Z}_1(A)^2\right\} = \mathsf{E}\left\{\mathcal{Z}_2(A)^2\right\} = 2\int_A \gamma(\omega)\,d\omega$$

for Borel sets $A \subset [0, \pi]$. Define a time series by

$$X_t = \int_0^\pi \cos(\omega t)\mathcal{Z}_1(d\omega) + \int_0^\pi \sin(\omega t)\mathcal{Z}_2(d\omega) \tag{12}$$

Then $\mathsf{E}X_t = 0$ for each t since each $\mathcal{Z}_i(\cdot)$ has mean zero, and the covariance is

$$\begin{aligned} C(s,t) &= 2\int_0^\pi \cos(\omega s)\cos(\omega t)\gamma(\omega)\,d\omega + 2\int_0^\pi \sin(\omega s)\sin(\omega t)\gamma(\omega)\,d\omega \\ &= 2\int_0^\pi \cos\big(\omega(s-t)\big)\,\gamma(\omega)\,d\omega, \end{aligned}$$

a function only of $h = (s - t)$, where $C(s,t) = c(s-t)$ with

$$c(h) = \oint e^{-i\omega h}\gamma(\omega)\,d\omega, \tag{13a}$$

where "$\oint$" denotes the integral over any interval of length 2π and where we have extended $\gamma(\omega)$ to be even on $[-\pi, \pi]$ and periodic on $\mathbb{R}$ (so the imaginary part of the integral vanishes). Thus X_t is a *stationary* Gaussian process, with mean zero. The function $\gamma(\omega)$, called the "spectral density", can be written as a Fourier series

$$\gamma(\omega) = \sum_{n=-\infty}^{\infty} \gamma_n e^{in\omega}$$

with real coefficients satisfying $2\pi\gamma_n = c(n) = 2\pi\gamma_{-n}$ by Fourier inversion using (13a). Thus

$$\begin{aligned} \gamma(\omega) &= \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} c(n)\, e^{in\omega} \\ &= \frac{1}{2\pi} c(0) + \frac{1}{\pi} \sum_{n=1}^{\infty} c(n)\, \cos(n\omega) \end{aligned} \tag{13b}$$

It turns out that *all* stationary time series have representation (12) and autocorrelation function with representation (13), or a slight generalization with $c(h) = \oint e^{-i\omega h} \gamma(d\omega)$ for an even positive Borel measure $\gamma(d\omega)$ on the circle.

For fixed constants $\{\omega_j\} \subset (0, \pi)$ and $\sigma_j^2 > 0$ and iid $\{U_j\}, \{V_j\} \overset{\text{iid}}{\sim} \mathsf{No}(0, 1)$ the stationary Gaussian process

$$X_t = \sum_j \sigma_j U_j \cos(\omega_j t) + \sum_j \sigma_j V_j \sin(\omega_j t)$$

has sample paths that oscillate at frequencies $\{\omega_j\}$, and has autocovariance

$$\begin{aligned} c(s,t) &= \sum_j \sigma_j^2 \cos\big(\omega_j (t - s)\big) \\ &= \oint e^{-i\omega(s-t)} \gamma(d\omega) \end{aligned}$$

with spectral measure

$$\gamma(d\omega) = \sum_j \sigma_j^2 [\delta_{\omega_j}(d\omega) + \delta_{-\omega_j}(d\omega)]/2$$

consisting of point masses of size σ_j^2 split evenly at pairs $\pm\omega_j$. Thus the spectral measure indicates how much "energy" $\gamma(A)$ the process has in frequency range A. For processes (like AR and MA processes) with continuous spectral densities, a peak for $\gamma(\omega)$ at $\omega^* \in (0, \pi)$ will be associated with oscillations at frequency $\pm\omega^*$.

3.1 Linear Filters

Let $\{X_t\}$ be a stationary mean-zero time series and construct a time series $\{Y_t\}$ as a linear combination

$$Y_t = \sum_{s=-\infty}^{\infty} a_k X_{t-k}$$

of $\{X_t\}$. This transformation is called a *linear filter*, and leads again to a stationary mean-zero time series with autocorrelation function

$$\begin{aligned}
c_Y(h) &= \mathsf{E} Y_t Y_{t+h} \\
&= \mathsf{E}\left\{ \sum_j a_j X_{t-j} \sum_k a_k X_{t+h-k} \right\} \\
&= \sum_{j,k} a_j a_k c_X(j-k+h) \\
&= \sum_{j,k} a_j a_k \oint e^{-i(j-k+h)\omega} \gamma_X(\omega)\, d\omega \\
&= \oint e^{-ih\omega}\, \overline{\Big\{ \sum_j a_j e^{ij\omega} \Big\}} \Big\{ \sum_k a_k e^{ik\omega} \Big\} \gamma_X(\omega)\, d\omega \\
&= \oint e^{-ih\omega} \Big| \sum_j a_j e^{ij\omega} \Big|^2 \gamma_X(\omega)\, d\omega \\
&= \oint e^{-i\omega h} \gamma_Y(\omega)\, d\omega,
\end{aligned}$$

from (13a), so the spectral densities for X_t and Y_t are related by

$$\gamma_Y(\omega) = \Big| \sum_j a_j e^{ij\omega} \Big|^2 \gamma_X(\omega). \tag{14}$$

Since an iid sequence $\{\zeta_t\} \sim \mathsf{No}(0, \sigma^2)$ has autocorrelation $c(h) = \sigma^2 \mathbf{1}_{\{h=0\}}$ and hence constant spectral density $\gamma(\omega) = \sigma^2/2\pi$, from (1a) and (14) we see that the spectral density for a Moving Average process is

$$\gamma_{\text{MA}}(\omega) = \frac{\sigma^2 |Q(e^{i\omega})|^2}{2\pi} \tag{15a}$$

for the polynomial $Q(z) = \sum_{i=0}^{q} b_i z^i$. Similarly, from (3) and (14), the spectral density for an Autoregressive process is

$$\gamma_{\text{AR}}(\omega) = \frac{\sigma^2}{2\pi |P(e^{i\omega})|^2} \tag{15b}$$

where $P(z) = \sum_{j=0}^{p} a_j\, z^j$.

4 Extensions: VAR and ARMA Processes

The n-dimensional form of Eqn (4),

$$Y(t) = RY(t-1) + Z_t$$

for $\mathbb{R}^n$-valued process $Y(t)$, iid $\mathbb{R}^n$-valued Z_t, and $n \times n$ matrix R with spectral radius less than one, is also Markov with simple exponential form for the cross-autocorrelation, and a convergent $\mathsf{MA}(\infty)$ representation. It is called a "vector auto-regressive process", denoted $Y(t) \sim \mathsf{VAR}(1)$. If $\{X_t\} \sim \mathsf{AR}(p)$ is real-valued, then the p-dimensional vector $Y_j(t) = (X_{t-j} :\ 0 \le j < p)$ is $\mathsf{VAR}(1)$ and satisfies the vector equation $Y(t) = RY(t-1) + Z_t$ for a particular matrix R:

$$\underbrace{\begin{bmatrix} X_t \\ X_{t-1} \\ X_{t-2} \\ \vdots \\ X_{t-p+2} \\ X_{t-p+1} \end{bmatrix}}_{Y(t)} = \underbrace{\begin{bmatrix} -a_1 & -a_2 & -a_3 & \cdots & -a_{p-1} & -a_p \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \dots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix}}_{R} \underbrace{\begin{bmatrix} X_{t-1} \\ X_{t-2} \\ X_{t-3} \\ \vdots \\ X_{t-p+1} \\ X_{t-p} \end{bmatrix}}_{Y(t-1)} + \underbrace{\begin{bmatrix} \zeta_t \\ 0 \\ 0 \\ \dots \\ 0 \\ 0 \end{bmatrix}}_{Z_t}$$

This gives an alternative approach to prediction and inference for $\mathsf{AR}(p)$ processes; for example, $\mathsf{E}[Y(t+h) \mid \mathcal{F}_t] = R^h Y(t)$, so $\mathsf{E}[X_{t+h} \mid \mathcal{F}_t] = e_1' R^h Y(t)$ where $e_1' = [1, 0, 0, \cdots, 0] \in \mathbb{R}^p$. Of course, everything now hinges on the spectral decomposition of the matrix R whose charactistic polynomial is

$$\chi(\lambda) = \sum_{j=0}^{p} a_j\, \lambda^{p-j} = \lambda^p P\big(\lambda^{-1}\big)$$

for the $P(z)$ from (7), and so whose eigenvalues are the inverses $\lambda_n = r_n^{-1}$ of the roots of $P(z)$.

4.1 ARMA Processes

The two concepts of $\mathsf{MA}(q)$ of Section (1) and $\mathsf{AR}(p)$ of Section (2) may be combined to construct a class of processes with the representation

$$\sum_{j=0}^{p} a_j X_{t-j} = \sum_{i=0}^{q} b_i \zeta_{t-i},$$

for $\zeta_t \overset{\text{iid}}{\sim} \mathsf{No}(0, \sigma^2)$, an AR process whose innovations are themselves a MA process. From Section (3.1) it follows that the spectral density function of $\{X_t\} \sim \mathsf{ARMA}(p, q)$ will be the rational function

$$\gamma(\omega) = \frac{\sigma^2 |Q(e^{i\omega})|^2}{2\pi |P(e^{i\omega})|^2}.$$

References

Walker, G. T. (1925), "On Periodicity (with disussion)," *Quaterly Journal of the Royal Meteorological Society*, 51, 337–346.

Yule, G. U. (1927), "On a method of investigating periodicities in disturbed series, with special reference to Wolfer's sunspot numbers," *Philosophical Transactions of the Royal Society: London Series A*, 226, 267–298.