Contents

1 Parametric spectral density estimation 3
  1.1 Statistical description of signals 3
  1.2 Linear Signal Models 6
  1.3 Memory of pole-zero models 10
  1.4 All-pole models 11
    1.4.1 By the way.... 11
    1.4.2 Impulse response 12
    1.4.3 Autocorrelation from predictor coefficients 13
    1.4.4 Autocorrelation from poles 16
    1.4.5 Yule-Walker equations 17
    1.4.6 Partial autocorrelations (PACS) 19
    1.4.7 Alternative parametrization strategies 20
  1.5 All-zero models 21
1 Parametric spectral density estimation

1.1 Statistical description of signals

Given a random variable $g$, expectation is

$$ E[g] = \mu(g) = \begin{cases} \sum_k g(k)p_k & g \text{ discrete} \\ \int_{-\infty}^{\infty} gp(g)dg & g \text{ continuous} \end{cases} $$

Expectation of the transformed random variable $f = \phi(g)$ is

$$ E[f] = E[\phi(g)] = \int_{-\infty}^{\infty} \phi(g)p(g)dg $$

Considering a signal sampled at $M$ discrete times as realization of an $M$-dimensional random variable:

**mean**

$$ \mu_g = E[g] = \begin{bmatrix} E[g_1] \\ \vdots \\ E[g_M] \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_M \end{bmatrix} $$

**(auto-) correlation matrix** with $r_{ij} = E[g_ig_j^*]$

$$ R_{gg} = E[gg^\dagger] = \begin{bmatrix} r_{11} & \cdots & r_{1M} \\ \vdots & \ddots & \vdots \\ r_{M1} & \cdots & r_{MM} \end{bmatrix} $$
**1 PARAMETRIC SPECTRAL DENSITY ESTIMATION**

(auto-) covariance matrix with $\sigma_{ij} = E[(g_i - \mu_i)(g_j - \mu_j)^{*}]$

\[ \Sigma_{gg} = E[(g - \mu_g)(g - \mu_g)^{\dagger}] = \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1M} \\ \vdots & \ddots & \vdots \\ \sigma_{M1} & \cdots & \sigma_{MM} \end{bmatrix} = R_{gg} - \mu_g\mu_g^{\dagger} \]

correlation coefficient

\[ \rho_{ij} = \frac{\sigma_{ij}}{\sigma_{ii}\sigma_{jj}} \]

reminder: Wiener-Khintchine theorem relates the power spectral density of a stationary stochastic process to the (auto-) correlation function

\[ S_{gg}(e^{j\omega}) = \sum_{l=-\infty}^{\infty} r_{gg}(l)e^{-j\omega l} \]
\[ r_{gg}(l) = \int_{-\infty}^{\infty} S_{gg}(e^{j\omega})e^{j\omega l}d\omega \]

**1.2 Linear Signal Models**

Most tractable are parametric models with rational system functions. Consider the constant coefficient difference equation with excitation / innovations / input $f$ and output $g$

\[ g(n) = -\sum_{k=1}^{P} a_k g(n-k) + \sum_{k=0}^{Q} d_k f(n-k) \]

From $g(n) = (h * f)(n) \leftrightarrow G(z) = H(z)F(z)$ and $G(z)\sum_{k=0}^{P} a_k z^{-k} = F(z)\sum_{k=0}^{Q} d_k z^{-k}$ it follows that

\[ H(z) = \left( \frac{d_0}{a_0} \right) \frac{\sum_{k=0}^{Q} d_k z^{-k}}{\sum_{k=0}^{P} a_k z^{-k}} = c_{\text{gain}} \frac{\prod_{k=1}^{Q} (1 - z_k z^{-1})}{\prod_{k=1}^{P} (1 - p_k z^{-1})} \]
Three causal systems:

- **Q = 0**: all-pole model \( AP(P) = AR(P) \) “autoregressive”
  \[
g(n) = -\sum_{k=1}^{P} a_k g(n - k)
\]

- **P = 0**: all-zero model \( AZ(Q) = MA(Q) \) “moving average”
  \[
g(n) = \sum_{k=0}^{Q} d_k f(n - k)
\]

- **P, Q > 0**: pole-zero model \( PZ(P, Q) = ARMA(P, Q) \)
  \[
g(n) = -\sum_{k=1}^{P} a_k g(n - k) + \sum_{k=0}^{Q} d_k f(n - k)
\]
If one of the three systems above is driven with white noise $f$ (independently and identically distributed realizations of a random variable) with zero mean and finite variance $\sigma_{ff}^2(0) = \sigma_w^2$, $\sigma_{ff}^2(l \neq 0) = 0$ then

$$r_{gg}(l) = \sigma_w^2 r_{hh}(l)$$
$$S_{gg}(e^{j\omega}) = \sigma_w^2|H(e^{j\omega})|^2 = \sigma_w^2 S_{hh}(e^{j\omega})$$
$$S_{gg}(z) = \sigma_w^2 H(z)H^*(1/z^*) = \sigma_w^2 S_{hh}(z)$$

### 1.3 Memory of pole-zero models

If $Q \leq P$, system function can be written (using polynomial division)

$$H(z) = \sum_{k=1}^{P} \frac{A_k}{1 - p_k z^{-1}}$$

else

$$H(z) = \sum_{j=0}^{Q-P} B_j z^{-j} + \sum_{k=1}^{P} \frac{A_k}{1 - p_k z^{-1}}$$

If $Q > P$ and the model is causal, the impulse response is given by contributions from the zeros, the real poles and complex conjugate poles

$$h(n) = \sum_{j=0}^{Q-P} B_j \delta(n - j) + \sum_{k=1}^{P_1} A_k p_k^n u(n) + \sum_{k=1}^{P_2} C_k r_k^n \cos(\omega_i n + \phi_i) u(n)$$

All stable PZ systems with one pole have an exponentially decaying impulse response with exponentially decaying autocorrelation ("short memory"). Rate of decay is determined by outermost pole.
1.4 All-pole models

are the most frequently used due to ease of analysis and parametrization

\[ H(z) = \frac{c_{\text{gain}}}{A(z)} = \frac{c_{\text{gain}}}{1 + \sum_{k=1}^{P} a_k z^{-k}} = \frac{c_{\text{gain}}}{\prod_{k=1}^{P} (1 - p_k z^{-1})} \]

1.4.1 By the way....

any causal finite-order, all-pole model can be substituted by an infinite-order all-zero model! Reasoning for a single pole (\(|a| < 1\)):

\[
H(z) = \sum_{k=0}^{\infty} h(k) z^{-k} = \frac{1}{1 - a/z} = \sum_{k=0}^{\infty} a^k z^{-k} = \lim_{N \to \infty} \sum_{k=0}^{N} a^k z^{-k}
\]

\[
H_N(z) = \sum_{k=0}^{N} a^k z^{-k} = \frac{1 - a^{N+1} / z^{N+1}}{1 - a/z}
\]

The latter has one pole at \(z = a\) and \(l = 1, \ldots, N\) zeros at \(z_l = ae^{2\pi j l}\).

Pole and zero \(z_0\) cancel \(\Rightarrow N \to \infty\) zeros remain, distributed on a circle through \(a\).

1.4.2 Impulse response

\[
H(z) + \sum_{k=1}^{P} a_k H(z) z^{-k} = c_{\text{gain}}
\]

\[
h(n) + \sum_{k=1}^{P} a_k h(n-k) = c_{\text{gain}} \delta(n)
\]

If the system is causal and stable, then all poles are inside the unit circle and the system is minimum-phase. Then

\[
h(n) = \begin{cases} 
0 & n < 0 \\
c_{\text{gain}} & n = 0 \\
- \sum_{k=1}^{P} a_k h(n-k) & n > 0 
\end{cases}
\]
Coefficients $a_k$ are also denoted “predictor coefficients” because they predict the impulse response exactly. Conversely,

$$a_n = -\frac{h(n)}{h(0)} - \sum_{k=1}^{n-1} a_k \frac{h(n-k)}{h(0)} \quad a_0 = 1, \; n > 0$$

The first values of the impulse response $h(n), 0 \leq n \leq P$ specify the all-pole filter completely.

### 1.4.3 Autocorrelation from predictor coefficients

The infinite impulse response of an all-pole system leads to an infinite autocorrelation sequence which cannot be written down explicitly. However, a recurrence relation in terms of the predictor coefficients can be obtained:

$$h(n) + \sum_{k=1}^{P} a_k h(n-k) = c_{\text{gain}} \delta(n)$$

$$h^*(n-l)h(n) + h^*(n-l) \sum_{k=1}^{P} a_k h(n-k) = h^*(n-l) c_{\text{gain}} \delta(n)$$

$$\sum_{n=-\infty}^{\infty} \sum_{k=0}^{P} a_k h^*(n-l) h(n-k) = c_{\text{gain}} \sum_{n=-\infty}^{\infty} h^*(n-l) \delta(n) \quad a_0 = 1$$

$$\sum_{k=0}^{P} a_k r_{hh}(l-k) = c_{\text{gain}} h^*(-l)$$

For $l > 0$, by causality,

$$\sum_{k=0}^{P} a_k r_{hh}(l-k) = 0$$

$$r_{hh}(l) = -\sum_{k=1}^{P} a_k r_{hh}(l-k) \quad l > 0$$

we have a recurrence relation
For $l = 0$, with $h(0) = c_{\text{gain}}$

$$\sum_{k=0}^{P} a_k r_{hh}(k) = |c_{\text{gain}}|^2$$

$$1 \cdot r_{hh}(0) + \sum_{k=1}^{P} a_k r_{hh}(k) = |c_{\text{gain}}|^2$$

$$r_{hh}(0) \left( r_{hh}(0) + \sum_{k=1}^{P} a_k r_{hh}(k) \right) = r_{hh}(0)|c_{\text{gain}}|^2$$

$$r_{hh}(0) = \frac{r_{hh}(0)|c_{\text{gain}}|^2}{r_{hh}(0) + \sum_{k=1}^{P} a_k r_{hh}(k)} = \frac{|c_{\text{gain}}|^2}{1 + \sum_{k=1}^{P} a_k \rho_{hh}(k)}$$

with the “normalized autocorrelation coefficient” $\rho(k) = r(k)/r(0)$.

Since $r_{hh}(0) = \int_{-\pi}^{\pi} S_{hh}(e^{j\omega})d\omega$ and $G(e^{j\omega}) = H(e^{j\omega})F(e^{j\omega})$ and $F(\delta) = 1$, this gives the energy of the output of an all-pole filter excited by a single impulse.

1.4.4 Autocorrelation from poles

The autocorrelation can also be obtained from the poles by the inverse $z$-transform:

$$S_{hh}(z) = H(z)H^*(1/z^*) = |c_{\text{gain}}|^2 \prod_{k=1}^{P} \frac{1}{(1 - p_k z^{-1})(1 - p_k z^*)}$$

$$r_{hh} = Z^{-1}(S_{hh})$$

Since the autocorrelation is symmetric (two-sided), so is the spectral density. The causal part of the system $H(z)$ contributes the r.h.s. and $H(1/z^*)$ the l.h.s.

Real poles contribute exponential decays, conjugate poles damped sinusoids.
1.4.5 Yule-Walker equations

allow computation of the $P + 1$ model parameters $(a_k, k \in \{1, P\}$ and $c_{\text{gain}}$) from the first (estimated) $P + 1$ autocorrelation coefficients by solving a set of $P + 1$ linear equations, $0 \leq l \leq P$:

$$
\sum_{k=0}^{P} a_k r_{hh}(l - k) = c_{\text{gain}} h^*(l) \quad a_0 = 1
$$

This approach is called \textbf{correlation matching} because the correspondence between model parameters and autocorrelation coefficients is one-to-one.

The solution can also be obtained from the autocorrelation of the output instead of the impulse response because with white-noise input $r_{gg}(l) = \sigma_w^2 r_{hh}(l)$.

Comments

- The above matrix is Hermitian and Toeplitz and its structure is exploited by efficient algorithms such as Levinson-Durbin.

- The above equations allow solution of the spectral factorization problem—determining the system function from the autocorrelation—with a set of \textit{linear} equations. This is a unique feature of all-pole models.

- As discussed above, if the all-pole system is causal and stable, there is a one-to-one mapping between model parameters and autocorrelation coefficients. An all-pole model can reproduce an arbitrary (!) set of autocorrelation coefficients.

- Problem: model size???

$\rightarrow$ PACS, AIC, BIC, diagnostic plots
1.4.6 Partial autocorrelations (PACS)

To compute the partial correlation between $g(i)$ and $g(i + l)$, eliminate from both $g(i)$ and $g(i + l)$ the part that can be predicted from $g(i + 1) \cdots g(i + l - 1)$ such that both $g(i)$ and $g(i + l)$ are uncorrelated with $g(i + 1) \cdots g(i + l - 1)$.

$$r_{gg}(l) = \text{Corr}(g(i), g(i + l))$$

$$r_{gg}^{\text{part}}(l) = \text{Corr}(g(i), g(i + l)|g(i + 1), \ldots, g(i + l - 1))$$

The partial autocorrelation sequence is found from the predictor coefficients $a_{P'}$ of AR($P'$) models of increasing orders $P' = 1, 2, \ldots$. They can thus be found by solution of multiple sets of YW-eqn’s.

By the central limit theorem, estimated partial autocorrelations of an AR($P$) process are approximately normally distributed with $E[r_{gg}^{\text{part}}] = 0$ and $\text{Var}(r_{gg}^{\text{part}}(l)) = 1/n$ for $l > P$.

$\longrightarrow$ 95% confidence interval is approximately $\pm 2/\sqrt{n}$.

1.4.7 Alternative parametrization strategies

**Burg’s algorithm**
minimizes sum of squares of forward and backward prediction error

**Least squares**
- forward ("covariance method")
- backward
- forward/backward: ("modified covariance method") – Percival & Walden’s favorite

**Maximum likelihood**
assume that time series is realization of a Gaussian process, find covariance matrix that maximizes that likelihood

Asymptotically, all of these are equal, but differences for small samples.
1.5 All-zero models

A causal AZ($Q$) model has the finite impulse response

$$h(k) = \begin{cases} 
d_k & 0 \leq k \leq Q \\
0 & \text{elsewhere}
\end{cases}$$

The autocorrelation of the impulse response is then

$$r_{hh}(l) = \sum_{k=0}^{\infty} h(k)h^*(k-l) = \begin{cases} 
\sum_{k=0}^{Q-l} d_k d_{k+l}^* & 0 \leq l \leq Q \\
0 & l > Q
\end{cases}$$

These $l$ equations are nonlinear and more difficult to solve.

The PACS of these models have an infinite extent since an invertible all-zero model is equivalent to an all-pole model of infinite order.

REFERENCES

Remembering our discussion on the “effect of single poles and zeros” on the magnitude of the amplitude response, AP systems are good at peak matching while AZ systems are good at trough matching.

References
