Aalto University

T-61.3040 Statistical Signal Modeling

Autumn 2011

Amaury Lendasse & Yoan Miche

Today's Topics (04.10)

- Power spectrum
- Filtering random processes
- Spectral factorization
- Power spectrum and generation of AR, MA and ARMA processes

Power Spectrum

- Related concepts: Spectral density, power spectral density (PSD), energy spectral density (ESD)
- Also called spectrum of the signal, sometimes
- Tells you how "the energy of the signal is distributed over frequency bins"
- The Energy spectral density is for signals of which you can take the Fourier transform of

Energy Spectral Density (ESD)

- If you have a signal $f_n, n \in \mathbb{Z}$, the Energy Spectral density $\Phi(\omega)$ is given by

$$\Phi(\omega) = \left| \frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{\infty} f_n e^{-i\omega n} \right|^2 = \frac{F(\omega)F^*(\omega)}{2\pi}$$

  with $F(\omega)$ the Discrete Time Fourier Transform of the signal f_n (assuming it exists)
- In our case, the $x(n)$ are random variables
- No guarantees on being integrable/summable (nor their square)
- Use the Power Spectral Density instead

Wiener-Khinchin Theorem and PSD

- The Fourier transform is not taken on the random process itself (not possible anyway), but on its autocorrelation function $r_x(k)$
- Again, we consider only WSS processes here
- The Wiener–Khinchin theorem states that the Power Spectral Density $P_x(e^{j\omega})$ is the Fourier transform of the autocorrelation function $r_x(k)$

$$P_x(e^{j\omega}) = \sum_{k=-\infty}^{\infty} r_x(k) e^{-jk\omega}$$

Wiener-Khinchin Theorem and PSD

- And the PSD and autocorrelation function form a Fourier transform pair, i.e.

$$r_x(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} P_x(e^{j\omega}) e^{jk\omega} d\omega$$

- So, PSD gives a frequency description of the second order moment of the process $x(n)$, that is, its "power"

Power Spectrum: Properties

- Three properties of the PSD:
  - Symmetry: If $x(n)$ is WSS, then the PSD is real-valued (i.e. $P_x(e^{j\omega}) = P_x^*(e^{j\omega})$) and $P_x(z)$ (z-transform version of the PSD) satisfies the symmetry

    $$P_x(z) = P_x^*(1/z^*).$$

    Also, if $x(n)$ is real-valued, $P_x(e^{j\omega}) = P_x(e^{-j\omega})$ (or $P_x(z) = P_x^*(z^*)$ for the z-transform version)
  - Positivity: The PSD of a WSS process is nonnegative: $P_x(e^{j\omega}) \geq 0$

Power Spectrum: Properties

- Three properties of the PSD (cont.):
  - Total Power: The power in a zero mean WSS process is proportional to the area under the PSD curve,

$$E\left[\left|x(n)\right|^2\right] = \frac{1}{2\pi}\int_{-\pi}^{\pi} P_x(e^{j\omega})d\omega$$

- Remember that $r_x(0) = E\left[\left|x(n)\right|^2\right]$, for this last property

Power Spectrum: White noise

- An example, with a white noise process
- What is "white noise"?
- A WSS process $v(n)$ is said to be *white* if the autocovariance is 0 for all $k \neq 0$, i.e.

$$c_v(k) = \sigma_v^2 \delta(k)$$

with $\delta(k)$ the unit sample

Power Spectrum: White noise

- White noise: sequence of uncorrelated random variables, each with variance σ_v^2
- E.g., sequence of uncorrelated Gaussian RVs is a white noise process

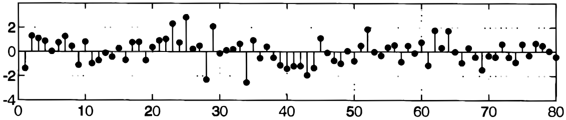

Figure: White Gaussian Noise (WGN). From Hayes book.

- Now, the PSD is then a constant: $P_v(e^{j\omega}) = \sigma_v^2$

Power Spectrum: Harmonic process

For the case of the random phase sinusoid $x(n) = A\sin(n\omega_0 + \phi)$, remember we have $r_x(k) = \frac{1}{2}A^2\cos(k\omega_0)$, which gives a PSD

$$P_x(e^{j\omega}) = \frac{1}{2}\pi A^2\left[u_0(\omega - \omega_0) + u_0(\omega + \omega_0)\right]$$

Power Spectrum and autocorrelation matrix

- A last property, on the eigenvalues of the autocorrelation matrix
- *Eigenvalue Extremal Property*: The eigenvalues of the autocorrelation matrix of a zero mean WSS process are bounded by the extremal values of the PSD, i.e.

$$\min_{\omega} P_x(e^{j\omega}) \leq \lambda_i \leq \max_{\omega} P_x(e^{j\omega})$$

Filtering

"Mapping an input to an output signal"

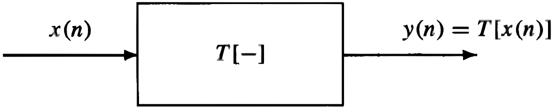

Figure: Filtering: a discrete time system that maps $x(n)$ to $y(n)$. From Hayes book.

Filtering: Linear Shift-Invariance property

- Linearity of a discrete-time system: with two inputs $x_1(n)$ and $x_2(n)$ and two constants a and b, we have

$$T\left[ax_1(n) + bx_2(n)\right] = aT\left[x_1(n)\right] + bT\left[x_2(n)\right]$$

- Which is important for example in the case where the input is a superposition of unit samples like

$$x(n) = \sum_{k=-\infty}^{\infty} x(k)\delta(n-k)$$

Filtering: Linear Shift-Invariance property

Then, we have the output (using the linearity property)

$$y(n) = \sum_{k=-\infty}^{\infty} x(k) T\left[\delta(n-k)\right] = \sum_{k=-\infty}^{\infty} x(k) h_k(n)$$

with $h_k(n) = T\left[\delta(n-k)\right]$ the response of the system to the delayed unit sample $\delta(n-k)$

Filtering: Linear Shift-Invariance property

- Shift-invariance of a discrete-time system: If a shift in the input results in the same shift in the output. That is, if we input $x(n - n_0)$, the output is $y(n - n_0)$.
- A system which is both linear and shift-invariant is called LSI (linear shift-invariant)

Filtering: Linear Shift-Invariance property

- So, if we have a WSS process $x(n)$ of mean m_x and autocorrelation $r_x(k)$
- We filter $x(n)$ with a stable LSI filter having a unit sample response $h(n)$
- Then the output $y(n)$ is a random process related to $x(n)$ by

$$y(n) = x(n) * h(n) = \sum_{k=-\infty}^{\infty} h(k)x(n-k)$$

Filtering: Linear Shift-Invariance property

- The mean of the output $y(n)$ can be found by

$$\begin{aligned} E\left[y(n)\right] &= E\left[\sum_{k=-\infty}^{\infty} h(k)x(n-k)\right] \\ &= \sum_{k=-\infty}^{\infty} h(k)E\left[x(n-k)\right] \\ &= m_x \sum_{k=-\infty}^{\infty} h(k) = m_x H(e^{j0}) \end{aligned}$$

- So the mean of the output is related to that of the input, scaled by the filter frequency response at $\omega = 0$

Filtering: Linear Shift-Invariance property

- In the same way, for autocorrelations of $y(n)$ and of $x(n)$, let's look at the cross-correlation first (needed to express $r_y(k)$)

$$
\begin{aligned}
r_{xy}(n+k,n) &= E\left[y(n+k)x^*(n)\right] \\
&= E\left[\sum_{l=-\infty}^{\infty} h(l)x(n+k-l)x^*(n)\right] \\
&= \sum_{l=-\infty}^{\infty} h(l)E\left[x(n+k-l)x^*(n)\right] \\
&= \sum_{l=-\infty}^{\infty} h(l)r_x(k-l)
\end{aligned}
$$

Filtering: Linear Shift-Invariance property

- So the cross-correlation $r_{xy}(n+k,n)$ only depends on k and we can write $r_{xy}(k) = r_x(k) * h(k)$
- And now

$$
\begin{aligned}
r_y(n+k,n) &= E\left[y(n+k)y^*(n)\right] \\
&= E\left[y(n+k)\sum_{l=-\infty}^{\infty} x^*(l)h^*(n-l)\right] \\
&= \sum_{l=-\infty}^{\infty} h^*(n-l)E\left[y(n+k)x^*(l)\right] \\
&= \sum_{l=-\infty}^{\infty} h^*(n-l)r_{xy}(n+k-l)
\end{aligned}
$$

Filtering: Linear Shift-Invariance property

- Which can be rewritten as

$$r_y(n+k, n) = \sum_{m=-\infty}^{\infty} h^*(m) r_{xy}(m+k) = r_{xy}(k) * h^*(-k)$$

where $r_y(n+k, n)$ only depends on k

Filtering: Linear Shift-Invariance property

- So,

$$r_y(k) = r_{xy}(k) * h^*(-k)$$

- And finally,

$$r_y(k) = r_x(k) * h(k) * h^*(-k) = r_x(k) * r_h(k)$$

  with $r_h(k)$ the autocorrelation of the unit sample response $h(n)$

- So, if $x(n)$ is WSS, then $y(n)$ is also WSS if $\sigma_y^2 < \infty$ (which requires the filter to be stable)
- Also, $x(n)$ and $y(n)$ will be jointly WSS

Filtering: Linear Shift-Invariance property

- Now if $h(n)$ is finite in length and zero outside $[0, N-1]$ on which $x(n)$ is defined, the variance (power) of $y(n)$, we can write

$$\sigma_y^2 = E\left[\left|y(n)\right|^2\right] = \mathbf{h}^H \mathbf{R}_x \mathbf{h}$$

- And using the relationship between autocorrelation, we have one between power spectrums:

$$P_y(e^{j\omega}) = P_x(e^{j\omega})\left|H(e^{j\omega})\right|^2$$

- So the power spectrum of a WSS process filtered by a LSI filter is that of the input signal scaled by the squared magnitude of the frequency response of the filter

Filtering: Example with white noise

An example with white noise again:

- Say we have $y(n)$ the random process created by filtering the white noise $w(n)$ with a LSI filter of system function

$$H(z) = \frac{1}{1 - 0.25z^{-1}}$$

- Say we also have $\sigma_w^2 = 1$ (remember it is constant for all n), then the power spectrum of $x(n)$ is

$$P_x(z) = \sigma_w^2 H(z) H(z^{-1}) = \frac{1}{(1 - 0.25z^{-1})(1 - 0.25z)}$$

Filtering: Example with white noise

- Using partial fraction expansion, we get

$$P_x(z) = \frac{16/15}{1 - 0.25z^{-1}} - \frac{16/15}{1 - 4z^{-1}}$$

- Which using inverse z-transform, gives the autocorrelation of $x(n)$ as

$$r_x(k) = \frac{16}{15}\left(\frac{1}{4}\right)^k u(k) + \frac{16}{15}4^k u(-k-1) = \frac{16}{15}\left(\frac{1}{4}\right)^{|k|}$$

Filtering: Example with white noise

Another example, where we start from a design constraint on the power spectrum

- Say we want to create a process whose power spectrum is of the form
$$P_x(e^{j\omega}) = \frac{5 + 4\cos 2\omega}{10 + 6\cos 2\omega}$$
- And we want to have it by filtering unit variance white noise with a LSI filter

Filtering: Example with white noise

- Let's expand $P_x(e^{j\omega})$:

$$P_x(e^{j\omega}) = \frac{5 + 2e^{2j\omega} + 2e^{-2j\omega}}{10 + 3e^{j\omega} + 3e^{-j\omega}} = \frac{(2z^2 + 1)(2z^{-2} + 1)}{(3z + 1)(3z^{-1} + 1)}$$

- And then factorize it in terms of system functions $P_x(z) = H(z)H(z^{-1})$, with

$$H(z) = \frac{2z^2 + 1}{3z + 1} = z\frac{2}{3}\frac{1 + \frac{1}{2}z^{-2}}{1 + \frac{1}{3}z^{-1}}$$

- Which is a stable filter (all poles inside unit circle), so the output will have the desired power spectrum

Spectral Factorization

- The power spectrum $P_x(e^{j\omega})$ of the WSS process $x(n)$ is:
  - real-valued
  - positive
  - periodic function of ω
- The spectral factorization of $P_x(e^{j\omega})$ aims at expressing it as

$$P_x(z) = \sigma_0^2 Q(z) Q^*(1/z^*)$$

Spectral Factorization

Such a process is called *regular* process and has specific properties:

- A regular process can be realized by filtering white noise of variance σ_0^2 by a causal (LSI and output only depends on present and past values) and stable filter. This realization is called *innovations representation* of the process
- The inverse of this causal and stable filter, $1/H(z)$, is a *whitening filter.* So if we filter the process $x(n)$ by it, we obtain a white noise of variance σ_0^2. The formation of this white noise process is called *innovations process*
- Since the white noise and the regular process are related by an invertible transformation, one can be obtained from the other "easily"

Spectral Factorization

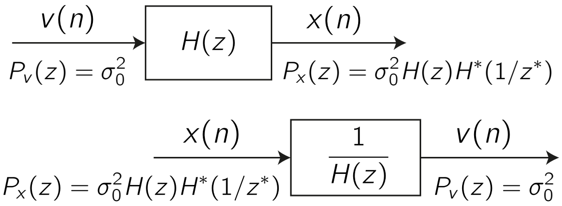

Figure: Innovations Representation and Innovations Process

Predictable process

- *Predictable process* (a.k.a. *deterministic*): a random process $x(n)$ such that there exists coefficients $a(k)$ to express $x(n)$ as

$$x(n) = \sum_{k=1}^{\infty} a(k)x(n-k)$$

  so $x(n)$ can be predicted without error by a linear combination of the previous values
- A process $x(n)$ is *predictable* iff its spectrum is made of impulses:

$$P_x(e^{j\omega}) = \sum_{k=1}^{N} \alpha_k u_0(\omega - \omega_k)$$

Wold Decomposition Theorem

Wold Decomposition Theorem: any WSS random process $x(n)$ can be written as the sum of two processes $x_{\text{pred}}(n)$ and $x_{\text{reg}}(n)$, where $x_{\text{pred}}(n)$ is a *predictable* process and $x_{\text{reg}}(n)$ a *regular* process, with $x_{\text{reg}}(n)$ and $x_{\text{pred}}(n)$ *orthogonal*, i.e.

$$E\left[x_{\text{reg}}(m)x^*_{\text{pred}}(n)\right] = 0$$

Power Spectrum of ARMA process

General case: ARMA

- Suppose we filter white noise $v(n)$ with a causal LSI filter which has a rational system function with p poles and q zeros:

$$H(z) = \frac{B_q(z)}{A_p(z)} = \frac{\sum_{k=0}^{q} b_q(k) z^{-k}}{1 + \sum_{k=1}^{p} a_p(k) z^{-k}}$$

- Now, if the filter $H(z)$ is stable, output process $x(n)$ is WSS and if $P_v(z) = \sigma_v^2$, we have the power spectrum

$$P_x(z) = \sigma_v^2 \frac{B_q(z) B_q^*(1/z^*)}{A_p(z) A_p^*(1/z^*)}$$

- This power spectrum defines an autoregressive moving average process of order (p, q), i.e. ARMA(p, q)

Power Spectrum of ARMA process

- Now, we know that $x(n)$ and $v(n)$ are related by

$$x(n) + \sum_{l=1}^{p} a_p(l)x(n-l) = \sum_{l=0}^{q} b_q(l)v(n-l)$$

- We can get the same relation between their autocorrelations (multiply by $x^*(n-k)$ and expectation):

$$r_x(k) + \sum_{l=1}^{p} a_p(l)r_x(k-l) = \sum_{l=0}^{q} b_q(l)r_{vx}(k-l)$$

Power Spectrum of ARMA process

- Developing the cross-correlation and substituting, we get the *Yule-Walker* equations for an ARMA(p, q) process:

$$r_x(k) + \sum_{l=1}^{p} a_p(l) r_x(k-l) = \begin{cases} \sigma_v^2 c_q(k) & , 0 \leq k \leq q \\ 0 & , k > q \end{cases}$$

  with $c_q(k) = \sum_{l=0}^{q-k} b_q(l+k) h^*(l)$
- Note that using these equations, one can extrapolate the autocorrelation sequence $r_x(k)$ from a finite set of values
- One can also estimate the filter coefficients $a_p(k)$ and $b_q(k)$, but since the Yule-Walker equations are non-linear in these coefficients, it might get difficult

Power Spectrum of ARMA process: Example

Example: Filter white noise with a LSI filter with zeros $z = 0.95e^{\pm j\pi/2}$ and poles $z = 0.5e^{\pm j2\pi/5}$

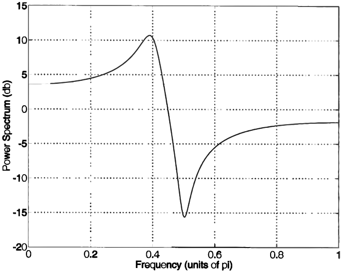

Figure: Power spectrum of an ARMA(2,2) process with zeros $z = 0.95e^{\pm j\pi/2}$ and poles $z = 0.5e^{\pm j2\pi/5}$.

Power Spectrum of AR process

- Autoregressive processes (AR(p))
- In this case, $x(n)$ is generated by filtering the white noise $v(n)$ by an all-pole filter

$$H(z) = \frac{b(0)}{1 + \sum_{k=1}^{p} a_p(k) z^{-k}}$$

- So, really, an AR process is an ARMA process with $q = 0$

Power Spectrum of AR process

- Again, if $P_v(z) = \sigma_v^2$, the power spectrum of $x(n)$ is

$$P_x(z) = \sigma_v^2 \frac{|b(0)|^2}{A_p(z)A_p^*(1/z^*)}$$

- And the Yule-Walker equations for an AR(p) process can be found by having $q = 0$ in the general case ARMA ones:

$$r_x(k) + \sum_{l=1}^{p} a_p(l) r_x(k-l) = \sigma_v^2 |b(0)|^2 \delta(k), k \geq 0$$

- Now, the Yule-Walker equations are linear in the filter coefficients $a_p(k)$, so they are easier to solve

Power Spectrum of AR process: Example

- Example: Given the first two autocorrelation values of a real-valued AR(1) process, assuming $\sigma_v^2 = 1$ and since we have the symmetry $r_x(k) = r_x(-k)$ for real processes

$$\begin{aligned} r_x(0) + \quad r_x(1)a(1) &= b^2(0) \\ r_x(0)a(1) &= -r_x(1) \end{aligned}$$

- So we get $a(1) = -\frac{r_x(1)}{r_x(0)}$ and $b(0) = \sqrt{\frac{r_x^2(0) - r_x^2(1)}{r_x(0)}}$ and we have determined the coefficients of the first order filter that generates an AR(1) process, with the given autocorrelation values
- Of course, one can do it the other way around: express the autocorrelation sequence using the filter coefficients

Power Spectrum of AR process: Plots

Examples of power spectrum plots for AR(1) processes:

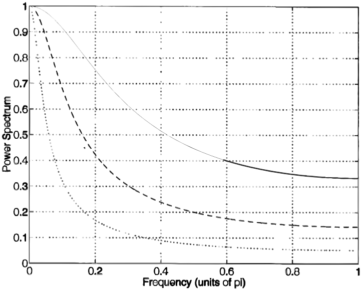

Figure: Power spectrum of a low-pass AR(1) process with a pole $z = 0.5$ (solid), $z = 0.75$ (dashed) and $z = 0.9$ (dotted).

Power Spectrum of AR process: Plots

Examples of power spectrum plots for AR(1) processes:

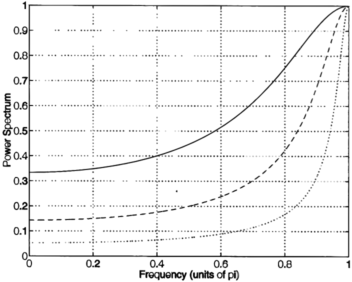

Figure: Power spectrum for a high-pass AR(1) process, with pole $z = -0.5$ (solid), $z = -0.75$ (dashed) and $z = -0.9$ (dotted).

Power Spectrum of MA process

Now, for MA processes

- MA(q) is ARMA(p, q) with $p = 0$
- So we generate $x(n)$ by filtering white noise with a filter (FIR) of the form

$$H(z) = \sum_{k=0}^{q} b_q(k) z^{-k}$$

- And if we have $P_v(z) = \sigma_v^2$, the power spectrum (recall the power spectrum of a WSS process filtered by a LSI filter, earlier) is

$$P_x(z) = \sigma_v^2 B_q(z) B_q^*(1/z^*)$$

Power Spectrum of MA process: Yule-Walker equations

- And the Yule-Walker equations can be found by using directly the inverse z-transform on the previous equation for the power spectrum

$$r_x(k) = \sigma_v^2 b_q(k) * b_q^*(-k) = \sigma_v^2 \sum_{l=0}^{q-|k|} b_q(l+|k|) b_q^*(l)$$

- So, the autocorrelation sequence of a MA(q) process is zero for k outside $[-q, q]$

Power Spectrum of MA process: Yule-Walker equations

- And again, estimating the MA(q) parameters from the autocorrelation is not easy, in general
- MA processes typically are slowly changing functions (w.r.t. frequency) with sharp nulls in the spectrum if $P_x(z)$ has zeros close to the unit circle

Power Spectrum of MA process

Example for a MA(4)

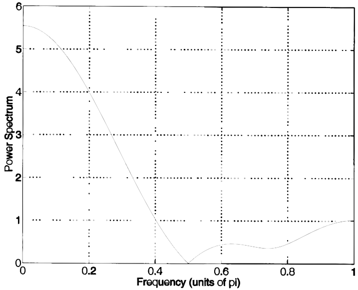

Figure: Power spectrum of an MA(4) process with zeros $z = e^{\pm j\pi/2}$ and $z = 0.8e^{\pm j3\pi/4}$. There are zeros on the unit circle at $\omega = \pm\pi/2$

Finally. . .

To summarize:

- mean, variance, autocorrelation, autocovariance of random processes
- stationarity, WSS, strict-sense stationarity, L-th order stationarity
- ergodicity (to use time averages as estimations of first/second order moments)
- power spectrum, spectral factorization
- whitening a WSS process or generate a WSS by filtering white noise
- using different types of filters (all-pole, zero-pole, FIR), generate AR, ARMA or MA, resp.

Next time

- Signal Modeling (aren't you here for that?)
- Levinson Recursion (maybe?)
- Lattice Filtering (likely the lecture after that)