Spectrum estimation using Periodogram, Bartlett and Welch

Guido Schuster

Slides follow closely chapter 8 in the book „Statistical Digital Signal Processing and Modeling“ by Monson H. Hayes and most of the figures and formulas are taken from there

Introduction

- We want to estimate the power spectral density of a wide-sense stationary random process

- Recall that the power spectrum is the Fourier transform of the autocorrelation sequence

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

- For an ergodic process the following holds

$$\lim_{N\to\infty} \left\{ \frac{1}{2N+1} \sum_{n=-N}^{N} x(n+k)x^*(n) \right\} = r_x(k)$$

Introduction

- The main problem of power spectrum estimation is
  - The data x(n) is always finite!

- Two basic approaches
  - Nonparametric (Periodogram, Bartlett and Welch)
    - These are the most common ones and will be presented in the next pages
  - Parametric approaches
    - not discussed here since they are less common

Nonparametric methods

- These are the most commonly used ones
- x(n) is only measured between n=0,..,N-1

$$\hat{r}_x(k) = \frac{1}{N}\sum_{n=0}^{N-1} x(n+k)x^*(n)$$

- Ensures that the values of x(n) that fall outside the interval [0,N-1] are excluded, where for negative values of k we use conjugate symmetry

$$\hat{r}_x(k) = \frac{1}{N}\sum_{n=0}^{N-1-k} x(n+k)x^*(n) \quad ; \quad k = 0, 1, \ldots, N-1$$

Periodogram

- Taking the Fourier transform of this autocorrelation estimate results in an estimate of the power spectrum, known as the Periodogram

$$\hat{P}_{per}(e^{j\omega}) = \sum_{k=-N+1}^{N-1} \hat{r}_x(k) e^{-jk\omega}$$

- This can also be directly expressed in terms of the data x(n) using the rectangular windowed function $x_N(n)$

$$x_N(n) = \begin{cases} x(n) & ; \quad 0 \le n < N \\ 0 & ; \quad \text{otherwise} \end{cases}$$

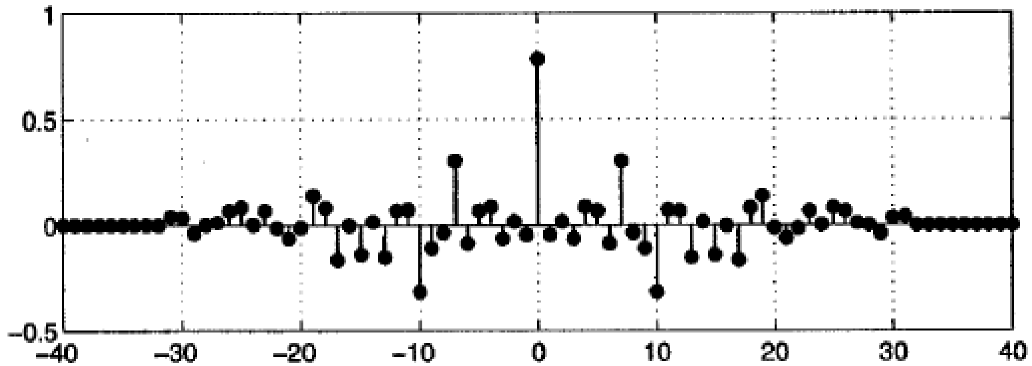
$$\hat{r}_x(k) = \frac{1}{N} \sum_{n=-\infty}^{\infty} x_N(n+k) x_N^*(n) = \frac{1}{N} x_N(k) * x_N^*(-k)$$

$$\boxed{\hat{P}_{per}(e^{j\omega}) = \frac{1}{N} X_N(e^{j\omega}) X_N^*(e^{j\omega}) = \frac{1}{N} \left| X_N(e^{j\omega}) \right|^2}$$

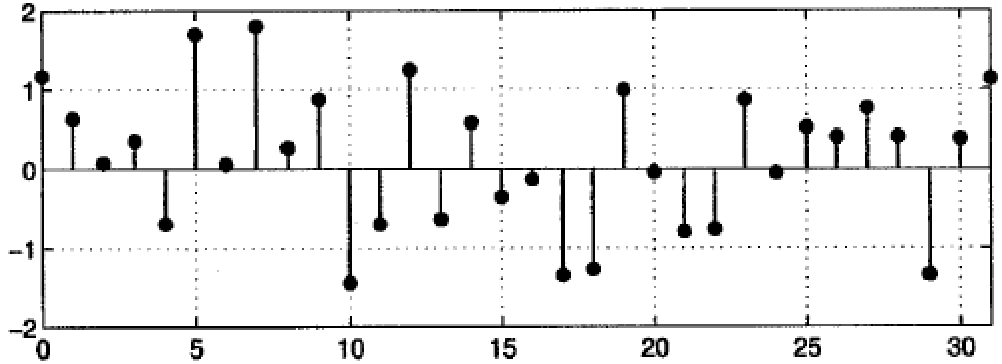
$$X_N(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x_N(n) e^{-jn\omega} = \sum_{n=0}^{N-1} x(n) e^{-jn\omega}$$

$$x_N(n) \xrightarrow{\text{DFT}} X_N(k) \longrightarrow \frac{1}{N} |X_N(k)|^2 = \hat{P}_{per}(e^{j2\pi k/N})$$

Periodogram of white noise

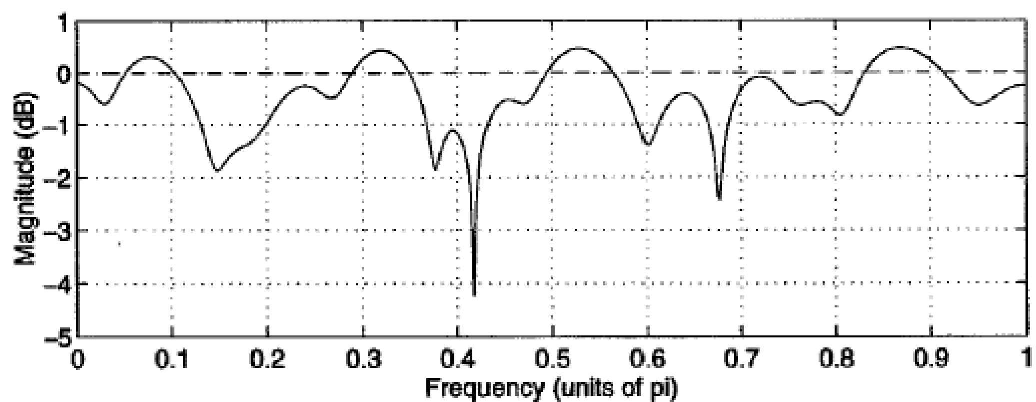


32 samples



$r_x(k) = \sigma_x^2 \delta(k)$

$P_x(e^{j\omega}) = \sigma_x^2$



Performance of the Periodogram

- If N goes to infinity, does the Periodogram converge towards the power spectrum in the mean squared sense?

$$\lim_{N\to\infty} E\left\{\left[\hat{P}_{per}(e^{j\omega}) - P_x(e^{j\omega})\right]^2\right\} = 0$$

- Necessary conditions
  - asymptotically unbiased:

$$\lim_{N\to\infty} E\left\{\hat{P}_{per}(e^{j\omega})\right\} = P_x(e^{j\omega})$$

  - variance goes to zero:

$$\lim_{N\to\infty} \mathrm{Var}\left\{\hat{P}_{per}(e^{j\omega})\right\} = 0$$

- In other words, it must be a consistent estimate of the power spectrum

Recall: sample mean as estimator

- Assume that we measure an iid process x[n] with mean μ and variance σ^2

- The sample mean is m=(x[0]+x[1]+x[2]+..+x[N-1])/N

- The sample mean is unbiased

  E[m] =E[(x[0]+x[1]+x[2]+..+x[N-1])/N]

  =(E[x[0]]+E[x[1]]+E[x[2]]+..+E[x[N-1]])/N

  $=N\mu/N$

  $=\mu$

- The variance of the sample mean is inversely proportional to the number of samples

  VAR[m] =VAR[(x[0]+x[1]+x[2]+..x[N-1])/N]=

  $=(VAR[x[0]]+VAR[x[1]]+VAR[x[2]]+..+VAR[x[N-1]])/N^2$

  $= N\sigma^2/N^2$

  $=\sigma^2/N$

Periodogram bias

- To compute the bias we first find the expected value of the autocorrelation estimate

$$E\{\hat{r}_x(k)\} = \frac{1}{N}\sum_{n=0}^{N-1-k} E\{x(n+k)x^*(n)\}$$

$$= \frac{1}{N}\sum_{n=0}^{N-1-k} r_x(k) = \frac{N-k}{N} r_x(k)$$

- Hence the estimate of the autocorrelation is biased with a triangular window (Bartlett)

$$E\{\hat{r}_x(k)\} = w_B(k) r_x(k)$$

$$w_B(k) = \begin{cases} \dfrac{N-|k|}{N} & ; \quad |k| \le N \\ 0 & ; \quad |k| > N \end{cases}$$

Periodogram bias

- The expected value of the Periodogram can now be calculated:

$$E\left\{\hat{P}_{per}(e^{j\omega})\right\} = E\left\{\sum_{k=-N+1}^{N-1} \hat{r}_x(k)e^{-jk\omega}\right\}$$

$$= \sum_{k=-N+1}^{N-1} E\left\{\hat{r}_x(k)\right\} e^{-jk\omega}$$

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

- Thus the expected value of the Periodogram is the convolution of the power spectrum with the Fourier transform of a Bartlett window

$$E\left\{\hat{P}_{per}(e^{j\omega})\right\} = \frac{1}{2\pi} P_x(e^{j\omega}) * W_B(e^{j\omega})$$

$$W_B(e^{j\omega}) = \frac{1}{N}\left[\frac{\sin(N\omega/2)}{\sin(\omega/2)}\right]^2$$

Periodogram bias

- Since the sinc-squared pulse converges towards a Dirac impulse as N goes to infinity, the Periodogram is asymptotically unbiased

$$\lim_{N\to\infty} E\left\{\hat{P}_{per}(e^{j\omega})\right\} = P_x(e^{j\omega})$$

Effect of lag window

- Consider a a random process consisting of a sinusoidal in white noise, where the phase of the sinusoidal is uniformly [-π, π] distributed

$$x(n) = A\sin(n\omega + \phi) + v(n)$$

- The power spectrum of such a signal is

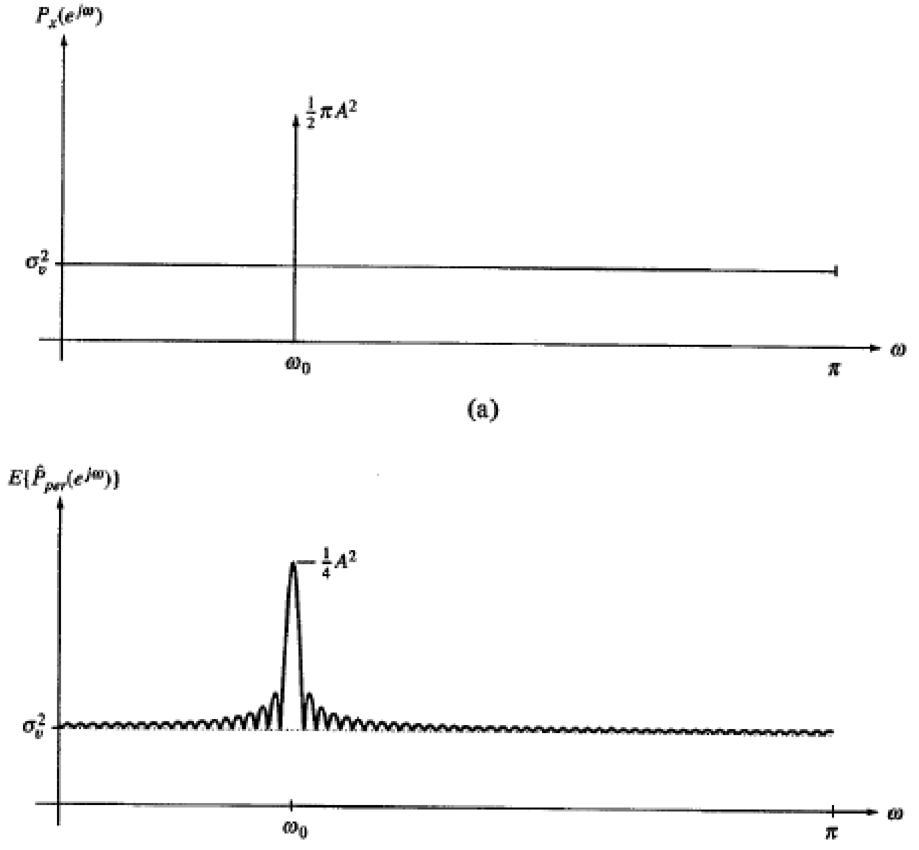
$$P_x(e^{j\omega}) = \sigma_v^2 + \frac{1}{2}\pi A^2\Big[u_0(\omega - \omega_0) + u_0(\omega + \omega_0)\Big]$$

- Therefore the expected value of the Periodogram is

$$E\left\{\hat{P}_{per}(e^{j\omega})\right\} = \frac{1}{2\pi}P_x(e^{j\omega}) * W_B(e^{j\omega})$$

$$= \sigma_v^2 + \frac{1}{4}A^2\left[W_B(e^{j(\omega-\omega_0)}) + W_B(e^{j(\omega+\omega_0)})\right]$$

Effect of lag window



Example: Periodogram of a Sinusoidal in Noise

- Consider a a random process consisting of a sinusoidal in white noise, where the phase of the sinusoidal is uniformly $[-\pi, \pi]$ distributed and $A=5$, $\omega_0=0.4\pi$
- N=64 on top and N=256 on the bottom
- Overlay of 50 Periodogram on the left and average on the right

$$x(n) = A\sin(n\omega_0 + \phi) + v(n)$$

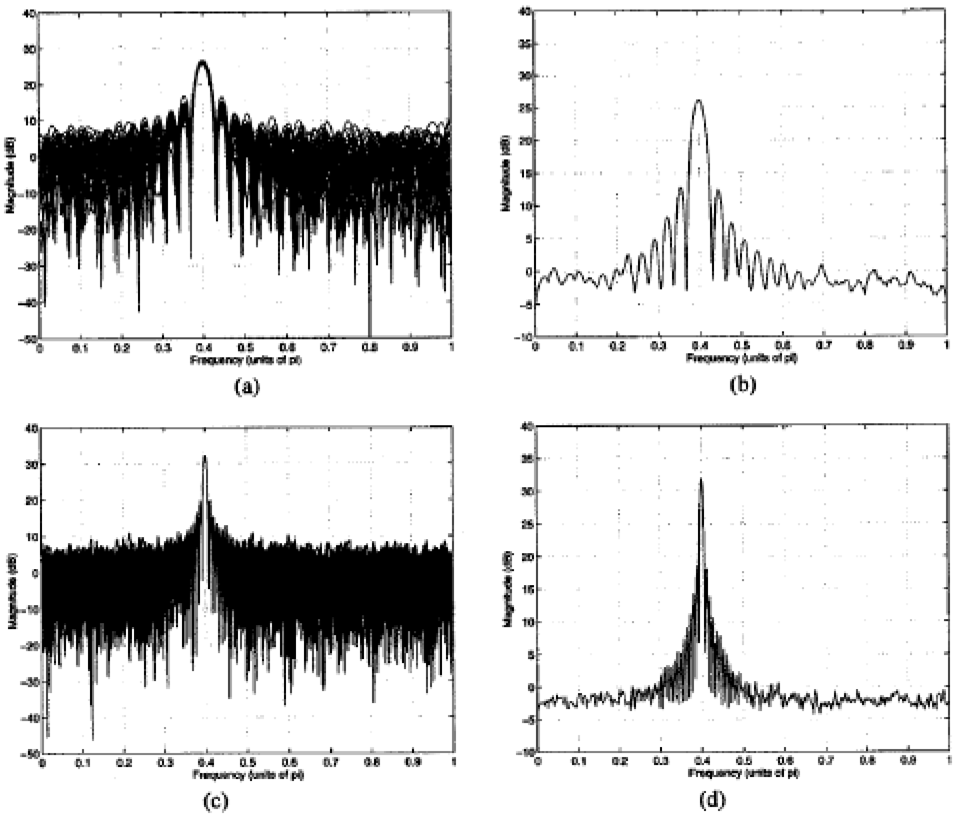


Figure 8.6 *The periodogram of a sinusoid in white noise. (a) Overlay plot of 50 periodograms using $N = 64$ data values and (b) the periodogram average. (c) Overlay plot of 50 periodograms using $N = 256$ data values and (d) the periodogram average.*

Periodogram resolution

- In addition to biasing the Periodogram, the spectral smoothing that is introduced by the Bartlett window also limits the ability of the Periodogram to resolve closely-spaced narrowband components

- Consider this random process consisting of two sinusoidal in white noise where the phases are again uniformly distributed and uncorrelated with each other

$$x(n) = A_1 \sin(n\omega_1 + \phi_1) + A_2 \sin(n\omega_2 + \phi_2) + v(n)$$

Periodogram resolution

- The power spectrum of the above random process is

$$P_x(e^{j\omega}) = \sigma_v^2 + \frac{1}{2}\pi A_1^2\Big[u_0(\omega - \omega_1) + u_0(\omega + \omega_1)\Big]$$
$$+ \frac{1}{2}\pi A_2^2\Big[u_0(\omega - \omega_2) + u_0(\omega + \omega_2)\Big]$$

- And the expected value of the Periodogram is

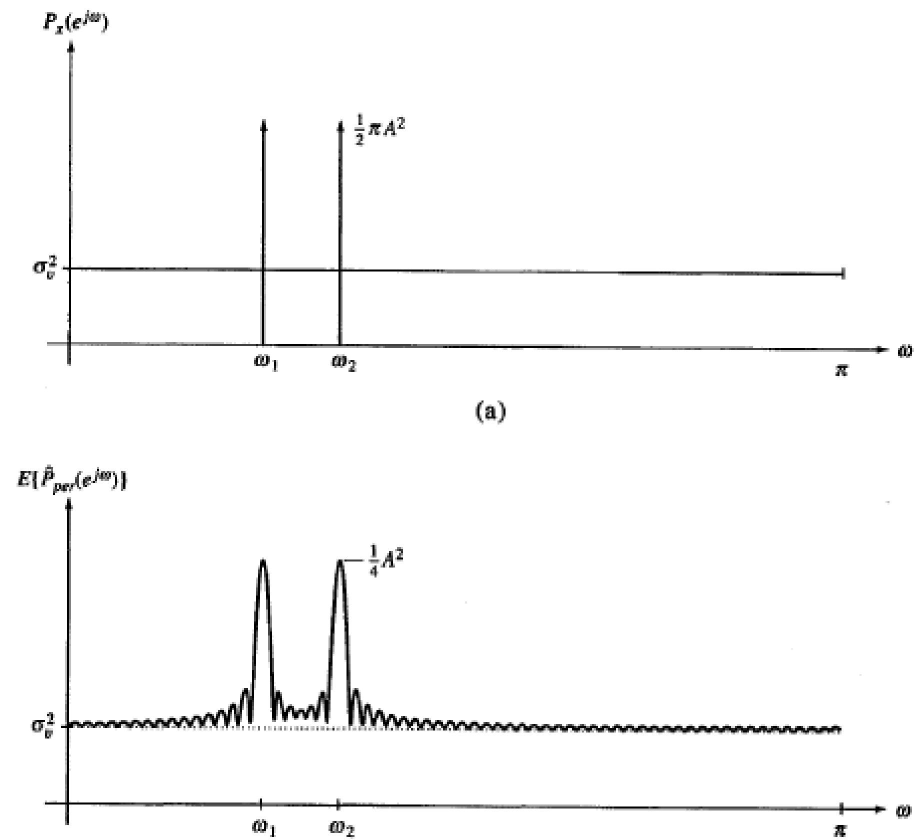
$$E\left\{\hat{P}_{per}(e^{j\omega})\right\} = \frac{1}{2\pi} P_x(e^{j\omega}) * W_B(e^{j\omega})$$
$$= \sigma_v^2 + \frac{1}{4}A_1^2\left[W_B(e^{j(\omega-\omega_1)}) + W_B(e^{j(\omega+\omega_1)})\right]$$
$$+\frac{1}{4}A_2^2\left[W_B(e^{j(\omega-\omega_2)}) + W_B(e^{j(\omega+\omega_2)})\right]$$

Periodogram resolution

- Since the width of the main lobe increases as N decreases, for a given N there is a limit on how closely two sinusoidal may be located before they can no longer be resolved

- This is usually defined as the bandwidth of the window at its half power points (-6dB), which is for the Bartlett window at 0.89*2π/N

- This is just a rule of thumb!



$$\text{Res}\left[\hat{P}_{per}(e^{j\omega})\right] = 0.89\frac{2\pi}{N}$$

Example: Periodogram of two Sinusoidal in Noise

- Consider a a random process consisting of two sinusoidal in white noise, where the phases of the sinusoidal are uniformly [-π, π] distributed and A=5, ω_1=0.4π , ω_2=0.45π
- N=40 on top
  and N=64 on the bottom
- Overlay of 50 Periodogram on the left and average on the right

$$x(n) = A\sin(n\omega_1 + \phi_1) + A\sin(n\omega_2 + \phi_2) + v(n)$$

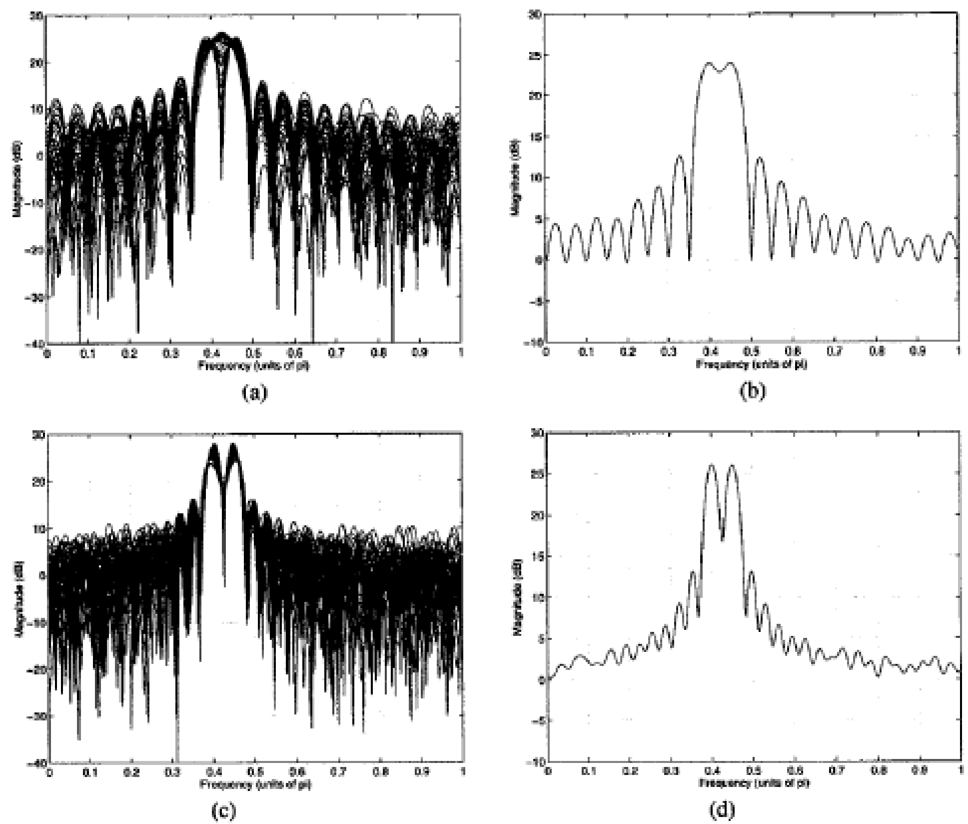


Figure 8.8 *The periodogram of two sinusoids in white noise with $\omega_1 = 0.4\pi$ and $\omega_2 = 0.45\pi$. (a) Overlay plot of 50 periodograms using $N = 40$ data values and (b) the ensemble average. (c) Overlay plot of 50 periodograms using $N = 64$ data values and (d) the ensemble average.*

Variance of the Periodogram

- The Periodogram is an asymptotically unbiased estimate of the power spectrum

$$\lim_{N\to\infty} E\left\{\hat{P}_{per}(e^{j\omega})\right\} = P_x(e^{j\omega})$$

- To be a consistent estimate, it is necessary that the variance goes to zero as N goes to infinity

$$\lim_{N\to\infty} \mathrm{Var}\left\{\hat{P}_{per}(e^{j\omega})\right\} = 0$$

- This is however hard to show in general and hence we focus on a white Gaussian noise, which is still hard, but can be done

Variance of the Periodogram

$$\hat{P}_{per}(e^{j\omega}) = \frac{1}{N}\left|\sum_{k=0}^{N-1} x(k)e^{-jk\omega}\right|^2 = \frac{1}{N}\left\{\sum_{k=0}^{N-1} x(k)e^{-jk\omega}\right\}\left\{\sum_{l=0}^{N-1} x^*(l)e^{jl\omega}\right\}$$

$$= \frac{1}{N}\sum_{k=0}^{N-1}\sum_{l=0}^{N-1} x(k)x^*(l)e^{-j(k-l)\omega} \tag{8.29}$$

Therefore, the second-order moment of the periodogram is

$$E\left\{\hat{P}_{per}(e^{j\omega_1})\hat{P}_{per}(e^{j\omega_2})\right\}$$

$$= \frac{1}{N^2}\sum_{k=0}^{N-1}\sum_{l=0}^{N-1}\sum_{m=0}^{N-1}\sum_{n=0}^{N-1} E\left\{x(k)x^*(l)x(m)x^*(n)\right\}e^{-j(k-l)\omega_1}e^{-j(m-n)\omega_2} \tag{8.30}$$

which depends on the fourth-order moments of $x(n)$. Since $x(n)$ is Gaussian, we may use the moment factoring theorem to simplify these moments [17,44]. For complex Gaussian random variables, the moment factoring theorem is[3]

$$\begin{aligned} E\left\{x(k)x^*(l)x(m)x^*(n)\right\} &= E\left\{x(k)x^*(l)\right\}E\left\{x(m)x^*(n)\right\} \\ &\quad + E\left\{x(k)x^*(n)\right\}E\left\{x(m)x^*(l)\right\} \end{aligned} \tag{8.31}$$

Variance of the Periodogram

Substituting Eq. (8.31) into Eq. (8.30), the second-order moment of the periodogram becomes a sum of two terms. The first term contains products of $E\{x(k)x^*(l)\}$ with $E\{x(m)x^*(n)\}$. For white noise, these terms are equal to σ_x^4 when $k = l$ and $m = n$, and they are equal to zero otherwise. Thus, the first term simplifies to

$$\frac{1}{N^2}\sum_{k=0}^{N-1}\sum_{m=0}^{N-1}\sigma_x^4 = \sigma_x^4 \tag{8.32}$$

The second term, on the other hand, contains products of $E\{x(k)x^*(n)\}$ with $E\{x(m)x^*(l)\}$. Again, for white noise, these terms are equal to σ_x^4 when $k = n$ and $l = m$, and they are equal to zero otherwise. Therefore, the second term becomes

$$\begin{aligned}
\frac{1}{N^2}\sum_{k=0}^{N-1}\sum_{l=0}^{N-1}\sigma_x^4 e^{-j(k-l)\omega_1}\; e^{j(k-l)\omega_2} &= \frac{\sigma_x^4}{N^2}\sum_{k=0}^{N-1} e^{-jk(\omega_1-\omega_2)}\sum_{l=0}^{N-1} e^{jl(\omega_1-\omega_2)} \\
&= \frac{\sigma_x^4}{N^2}\left[\frac{1-e^{-jN(\omega_1-\omega_2)}}{1-e^{-j(\omega_1-\omega_2)}}\right]\left[\frac{1-e^{jN(\omega_1-\omega_2)}}{1-e^{j(\omega_1-\omega_2)}}\right] \\
&= \sigma_x^4\left[\frac{\sin N(\omega_1-\omega_2)/2}{N\sin(\omega_1-\omega_2)/2}\right]^2
\end{aligned} \tag{8.33}$$

Variance of the Periodogram

Combining Eq. (8.32) and Eq. (8.33) it follows that

$$E\left\{\hat{P}_{per}(e^{j\omega_1})\hat{P}_{per}(e^{j\omega_2})\right\} = \sigma_x^4\left\{1+\left[\frac{\sin N(\omega_1-\omega_2)/2}{N\sin(\omega_1-\omega_2)/2}\right]^2\right\} \tag{8.34}$$

Since

$$\mathrm{Cov}\left\{\hat{P}_{per}(e^{j\omega_1})\hat{P}_{per}(e^{j\omega_2})\right\} = E\left\{\hat{P}_{per}(e^{j\omega_1})\hat{P}_{per}(e^{j\omega_2})\right\} - E\left\{\hat{P}_{per}(e^{j\omega_1})\right\}E\left\{\hat{P}_{per}(e^{j\omega_2})\right\}$$

and $E\left\{\hat{P}_{per}(e^{j\omega})\right\} = \sigma_x^2$, then the covariance of the periodogram is

$$\mathrm{Cov}\left\{\hat{P}_{per}(e^{j\omega_1})\hat{P}_{per}(e^{j\omega_2})\right\} = \sigma_x^4\left[\frac{\sin N(\omega_1-\omega_2)/2}{N\sin(\omega_1-\omega_2)/2}\right]^2 \tag{8.35}$$

Finally, setting $\omega_1 = \omega_2$ we have, for the variance,

$$\mathrm{Var}\left\{\hat{P}_{per}(e^{j\omega})\right\} = \sigma_x^4 \tag{8.36}$$

Thus, the variance does not go to zero as $N \rightarrow \infty$, and the periodogram is *not a consistent estimate* of the power spectrum. In fact, since $P_x(e^{j\omega}) = \sigma_x^2$ then the variance of the periodogram of white Gaussian noise is proportional to the square of the power spectrum,

$$\mathrm{Var}\left\{\hat{P}_{per}(e^{j\omega})\right\} = P_x^2(e^{j\omega}) \tag{8.37}$$

Example: Periodogram of white Gaussian noise

- For a white Gaussian noise with variance 1, the following holds

$$P_x(e^{j\omega}) = 1$$

- The expected value of the Periodogram is

$$E\left\{\hat{P}_{per}(e^{j\omega})\right\} = \frac{1}{2\pi} P_x(e^{j\omega}) * W_B(e^{j\omega})$$

- Which results in

$$E\{\hat{P}_{per}(e^{j\omega})\} = 1$$

- And the variance is

$$\text{Var}\left\{\hat{P}_{per}(e^{j\omega})\right\} = P_x^2(e^{j\omega})$$

- Which results in

$$\text{Var}\left\{\hat{P}_{per}(e^{j\omega})\right\} = 1$$

Example: Periodogram of white Gaussian noise

- On the left the overlay of 50 Periodograms are shown and on the right the average
- From top to bottom the data record length N increases from 64 to 128 to 256
- Note that the variance of the power spectrum estimate does not decrease, when N increases!

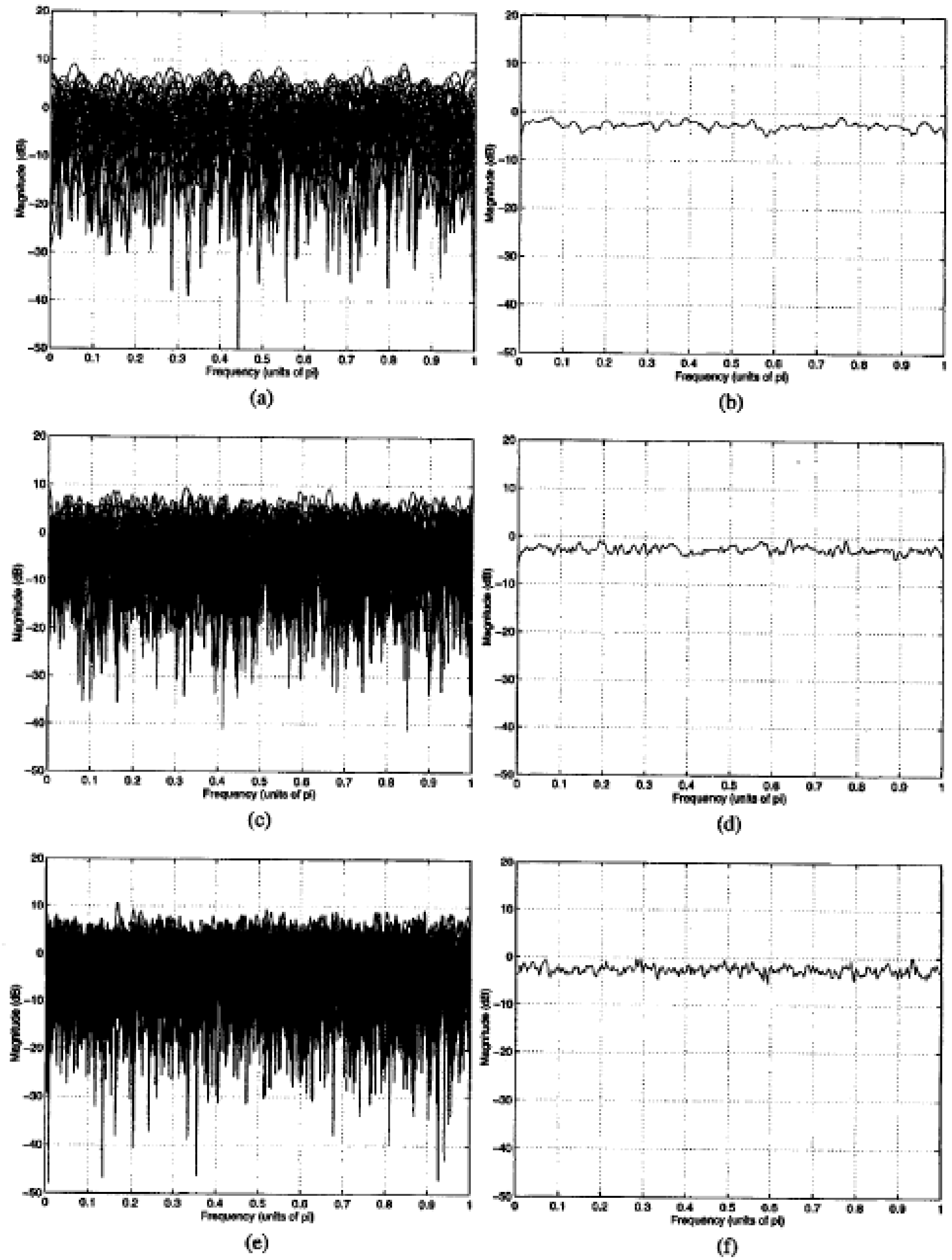


Figure 8.9 *The periodogram of unit variance white Gaussian noise. (a) Overlay plot of 50 periodograms with $N = 64$ data values and (b) the periodogram average. (c) Overlay plot of 50 periodograms with $N = 128$ data values and (d) the periodogram average. (e) Overlay plot of 50 periodograms with $N = 256$ data values and (f) the periodogram average.*

So what if the process is not white and/or not Gaussian?

- Interpret the process as filtered white noise v(n) with unit variance

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

- The white noise process and the colored noise process have the following Periodograms

$$\hat{P}_{per}^{(v)}(e^{j\omega}) = \frac{1}{N}\left|V_N(e^{j\omega})\right|^2$$

$$\hat{P}_{per}^{(x)}(e^{j\omega}) = \frac{1}{N}\left|X_N(e^{j\omega})\right|^2$$

- Although $x_N(n)$ is NOT equal to the convolution of $v_N(n)$ and h(n), if N is large compared to the length of h(n) then the transient effects are small

$$x_N(n) \approx h(n) * v_N(n)$$

So what if the process is not white and/or not Gaussian?

- With this interpretation, the result is, that the variance of the spectrum estimate is approximately the square of the true power spectrum
- Note that this is a function of ω
- Note that this is not good
  - Not a consistent estimate

$$x_N(n) \approx h(n) * v_N(n) \qquad |H(e^{j\omega})|^2 = P_x(e^{j\omega})$$

$$|X_N(e^{j\omega})|^2 \approx |H(e^{j\omega})|^2 \, |V_N(e^{j\omega})|^2 = P_x(e^{j\omega}) \, |V_N(e^{j\omega})|^2$$

$$\hat{P}^{(x)}_{per}(e^{j\omega}) = \frac{1}{N} |X_N(e^{j\omega})|^2 \qquad \hat{P}^{(v)}_{per}(e^{j\omega}) = \frac{1}{N} |V_N(e^{j\omega})|^2$$

$$\hat{P}^{(x)}_{per}(e^{j\omega}) \approx P_x(e^{j\omega}) \hat{P}^{(v)}_{per}(e^{j\omega})$$

1

$$\mathrm{Var}\left\{\hat{P}^{(x)}_{per}(e^{j\omega})\right\} \approx P_x^2(e^{j\omega}) \mathrm{Var}\left\{\hat{P}^{(v)}_{per}(e^{j\omega})\right\}$$

$$\mathrm{Var}\left\{\hat{P}^{(x)}_{per}(e^{j\omega})\right\} \approx P_x^2(e^{j\omega})$$

Periodogram summary

Table 8.1 *Properties of the Periodogram*

$$\hat{P}_{per}(e^{j\omega}) = \frac{1}{N}\left|\sum_{n=0}^{N-1} x(n)e^{-jn\omega}\right|^2$$

Bias

$$E\left\{\hat{P}_{per}(e^{j\omega})\right\} = \frac{1}{2\pi} P_x(e^{j\omega}) * W_B(e^{j\omega})$$

Resolution

$$\Delta\omega = 0.89\frac{2\pi}{N}$$

Variance

$$\text{Var}\left\{\hat{P}_{per}(e^{j\omega})\right\} \approx P_x^2(e^{j\omega})$$

The modified Periodogram

- What happens when another window (instead of the rectangular window) is used?
- The window shows itself in the Bias, but not directly but as a convolution with itself

$$\hat{P}_{per}(e^{j\omega}) = \frac{1}{N}|X_N(e^{j\omega})|^2 = \frac{1}{N}\left|\sum_{n=-\infty}^{\infty} x(n)w_R(n)e^{-jn\omega}\right|^2$$

$$E\left\{\hat{P}_{per}(e^{j\omega})\right\} =$$

$$= \frac{1}{N}E\left\{\left[\sum_{n=-\infty}^{\infty} x(n)w_R(n)e^{-jn\omega}\right]\left[\sum_{m=-\infty}^{\infty} x(m)w_R(m)e^{-jm\omega}\right]^*\right\}$$

$$= \frac{1}{N}E\left\{\sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty} x(n)x^*(m)w_R(m)w_R(n)e^{-j(n-m)\omega}\right\}$$

$$= \frac{1}{N}\sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty} r_x(n-m)w_R(m)w_R(n)e^{-j(n-m)\omega}$$

The modified Periodogram

- With the change of variables k=n-m this becomes

$$E\left\{\hat{P}_{per}(e^{j\omega})\right\} =$$

$$= \frac{1}{N}\sum_{k=-\infty}^{\infty}\sum_{n=-\infty}^{\infty} r_x(k)w_R(n)w_R(n-k)e^{-jk\omega}$$

$$= \frac{1}{N}\sum_{k=-\infty}^{\infty} r_x(k)\left[\sum_{n=-\infty}^{\infty} w_R(n)w_R(n-k)\right]e^{-jk\omega}$$

$$= \frac{1}{N}\sum_{k=-\infty}^{\infty} r_x(k)w_B(k)e^{-jk\omega}$$

- Where $w_B(k)$ is a Bartlett window

$$w_B(k) = w_R(k) * w_R(-k) = \sum_{n=-\infty}^{\infty} w_R(n)w_R(n-k)$$

- Hence in the frequency domain, this becomes

$$E\left\{\hat{P}_{per}(e^{j\omega})\right\} = \frac{1}{2\pi N}P_x(e^{j\omega}) * |W_R(e^{j\omega})|^2$$

$$W_R(e^{j\omega}) = \frac{\sin(N\omega/2)}{\sin(\omega/2)}e^{-j(N-1)\omega/2}$$

The modified Periodogram

- Smoothing is determined by the window that is applied to the data
- While the rectangular window as the smallest main lobe of all windows, its sidelobes fall off rather slowly

$$x(n) = 0.1\sin(n\omega_1 + \phi_1) + \sin(n\omega_2 + \phi_2) + v(n)$$

Rectangular data window Hamming data window

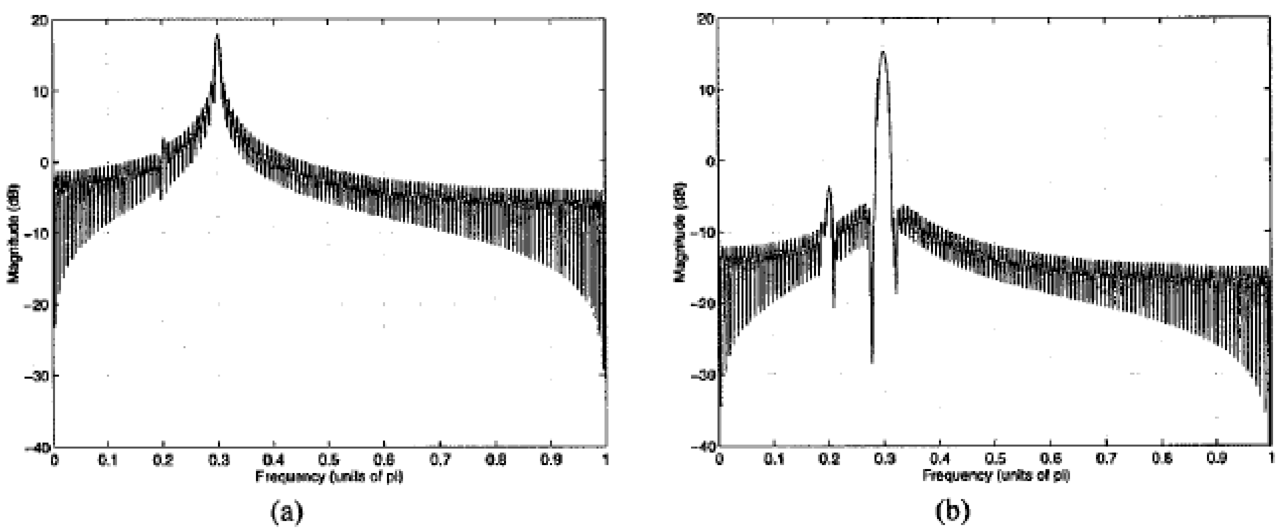


Figure 8.10 *Spectral analysis of two sinusoids in white noise with sinusoidal frequencies of* $\omega_1 = .2\pi$ *and* $\omega_2 = .3\pi$ *and a data record length of* $N = 128$ *points. (a) The expected value of the periodogram. (b) The expected value of the modified periodogram using a Hamming data window.*

The modified Periodogram

- Nothing is free. As you notice, the Hamming window has a wider main lobe

- The Periodogram of a process that is windowed with a general window is called modified Periodogram

- N is the length of the window and U is a constant that is needed so that the modified Periodogram is asymptotically unbiased

$$\hat{P}_M(e^{j\omega}) = \frac{1}{NU}\left|\sum_{n=-\infty}^{\infty} x(n)w(n)e^{-jn\omega}\right|^2$$

$$U = \frac{1}{N}\sum_{n=0}^{N-1} |w(n)|^2$$

The modified Periodogram

- For evaluating the Bias we take the expected value of the modified Periodogram, where $W(e^{j\omega})$ is the Fourier transform of the data window

$$E\left\{\hat{P}_M(e^{j\omega})\right\} = \frac{1}{2\pi NU} P_x(e^{j\omega}) * |W(e^{j\omega})|^2$$

- Using the Parseval theorem, it follows that U is the energy of the window divided by N

$$U = \frac{1}{N}\sum_{n=0}^{N-1}|w(n)|^2 = \frac{1}{2\pi N}\int_{-\pi}^{\pi}|W(e^{j\omega})|^2 d\omega$$

$$\frac{1}{2\pi NU}\int_{-\pi}^{\pi}|W(e^{j\omega})|^2 d\omega = 1$$

- With an appropriate window, $|W(e^{j\omega})|^2/NU$ will converge to an impulse of unit area and hence the modified Periodogram will be asymptotically unbiased

Variance of the modified Periodogram

- Since the modified Periodogram is simply the Periodogram of a windowed data sequence, not much changes

$$\mathrm{Var}\left\{\hat{P}_M(e^{j\omega})\right\} \approx P_x^2(e^{j\omega})$$

- Hence the estimate is still not consistent

- Main advantage is that the window allows a tradeoff between spectral resolution (main lobe width) and spectral masking (sidelobe amplitude)

Resolution versus masking of the modified Periodogram

- The resolution of the modified Periodogram defined to be the 3dB bandwidth of the data window

$$\text{Res}\left[\hat{P}_M(e^{j\omega})\right] = (\Delta\omega)_{3\text{dB}}$$

- Note that when we used the Bartlett lag window before, the resolution was defined as the 6dB bandwidth. This is consistent with the above definition, since the 3dB points of the data window transform into 6dB points in the Periodogram

Table 8.2 ***Properties of a Few Commonly Used Windows. Each Window is Assumed to be of Length N.***

Window	Sidelobe Level (dB)	3 dB BW $(\Delta\omega)_{3\text{dB}}$
Rectangular	−13	$0.89(2\pi/N)$
Bartlett	−27	$1.28(2\pi/N)$
Hanning	−32	$1.44(2\pi/N)$
Hamming	−43	$1.30(2\pi/N)$
Blackman	−58	$1.68(2\pi/N)$

Modified periodogram summary

Table 8.3 ***Properties of the Modified Periodogram***

	$\hat{P}_M(e^{j\omega}) = \frac{1}{NU} \left\| \sum_{n=-\infty}^{\infty} w(n)x(n)e^{-jn\omega} \right\|^2$
	$U = \frac{1}{N} \sum_{n=0}^{N-1} \|w(n)\|^2$
Bias	$E\left\{\hat{P}_M(e^{j\omega})\right\} = \frac{1}{2\pi NU} P_x(e^{j\omega}) * \|W(e^{j\omega})\|^2$
Resolution	Window dependent
Variance	$\text{Var}\left\{\hat{P}_M(e^{j\omega})\right\} \approx P_x^2(e^{j\omega})$

Bartlett's method

- Still have not a consistent estimate of the power spectrum!
- Nevertheless, the periodogram is asymptotically unbiased

$$\lim_{N\to\infty} E\{\hat{P}_{per}(e^{j\omega})\} = P_x(e^{j\omega})$$

- Hence if we can find a consistent estimate of the mean, then this estimate would also be a consistent estimate of the power spectrum

Bartlett's method

- Averaging (sampe mean) a set of uncorrelated measurements of a random variable results in a consistent estimate of its mean
- In other words: Variance of the sample mean is inversely proportional to the number of measurements
- Hence this should also work here, by averaging Periodograms

This suggests that we consider estimating the power spectrum of a random process by periodogram averaging. Thus, let $x_i(n)$ for $i = 1, 2, \ldots, K$ be K uncorrelated realizations of a random process $x(n)$ over the interval $0 \le n < L$. With $\hat{P}_{per}^{(i)}(e^{j\omega})$ the periodogram of $x_i(n)$,

$$\hat{P}_{per}^{(i)}(e^{j\omega}) = \frac{1}{L}\left|\sum_{n=0}^{L-1} x_i(n)e^{-jn\omega}\right|^2 \quad ; \quad i = 1, 2, \ldots, K$$

Bartlett's method

- Averaging these Periodograms

$$\hat{P}_x(e^{j\omega}) = \frac{1}{K}\sum_{i=1}^{K}\hat{P}_{per}^{(i)}(e^{j\omega})$$

- This results in an asymptotically unbiased estimate of the power spectrum

$$E\left\{\hat{P}_x(e^{j\omega})\right\} = E\left\{\hat{P}_{per}^{(i)}(e^{j\omega})\right\} = \frac{1}{2\pi}P_x(e^{j\omega}) * W_B(e^{j\omega})$$

- Since we assume that the realizations are uncorrelated, it follows, that the variance is inversely proportional to the number of measurements K

$$\mathrm{Var}\left\{\hat{P}_x(e^{j\omega})\right\} = \frac{1}{K}\mathrm{Var}\left\{\hat{P}_{per}^{(i)}(e^{j\omega})\right\} \approx \frac{1}{K}P_x^2(e^{j\omega})$$

- Hence this is a consistent estimate of the power spectrum, if L and K go to infinity

Bartlett's method

- There is still a problem: we usually do not have uncorrelated data records!
- Typically there is only one data record of length N available
- Hence Bartlett proposes to partition the data record into K nonoverlapping sequences of the length L, where N=K*L

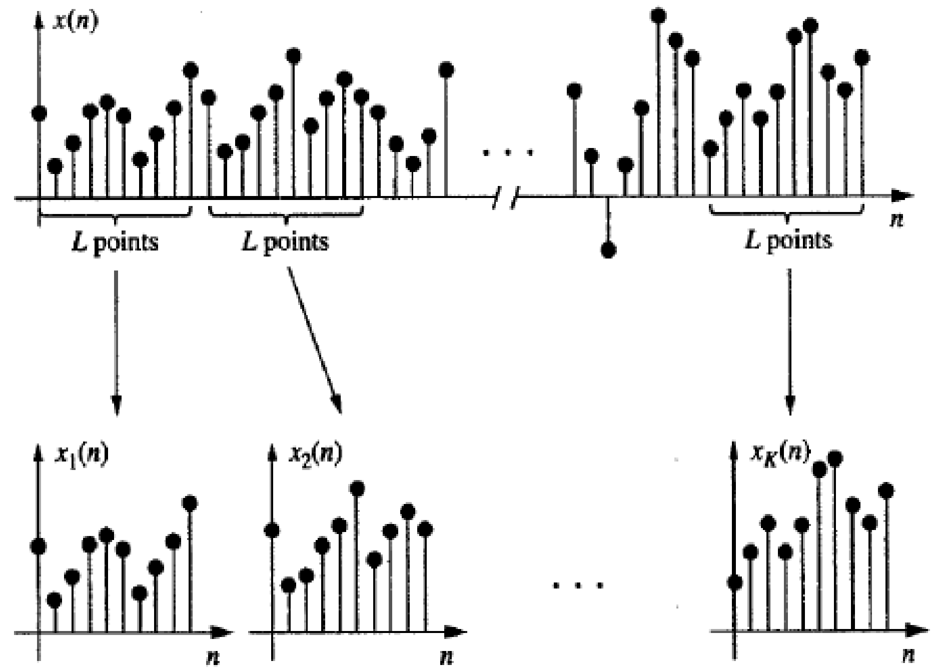


Figure 8.12 *Partitioning x(n) into nonoverlapping subsequences.*

$$x_i(n) = x(n+iL) \qquad \begin{array}{l} n = 0, 1, \ldots, L-1 \\ i = 0, 1, \ldots, K-1 \end{array}$$

Thus, the Bartlett estimate is

$$\boxed{\hat{P}_B(e^{j\omega}) = \frac{1}{N}\sum_{i=0}^{K-1}\left|\sum_{n=0}^{L-1} x(n+iL)e^{-jn\omega}\right|^2}$$

Bartlett’s method

- Each expected value of the periodogram of the subsequences are identical hence the process of averaging subsequences Periodograms results in the same average value => asymptotically unbiased

$$E\left\{\hat{P}_B(e^{j\omega})\right\} = \frac{1}{2\pi} P_x(e^{j\omega}) * W_B(e^{j\omega})$$

- Note that the data length used for the Periodograms are now L and not N anymore, the spectral resolution becomes worse (this is the price we are paying)

$$\text{Res}\left[\hat{P}_B(e^{j\omega})\right] = 0.89\frac{2\pi}{L} = 0.89K\frac{2\pi}{N}$$

Bartlett's method

- Now we reap the reward: the variance is going to zero as the number of subsequences goes to infinity

$$\text{Var}\left\{\hat{P}_B(e^{j\omega})\right\} \approx \frac{1}{K}\text{Var}\left\{\hat{P}_{per}^{(i)}(e^{j\omega})\right\} \approx \frac{1}{K}P_x^2(e^{j\omega})$$

- If both, K and L go to infinity, this will be a consistent estimate of the power spectrum

- In addition, for a given N=K*L, we can trade off between good spectral resolution (large L) and reduction in variance (Large K)

Table 8.4 *Properties of Bartlett's Method*

$$\hat{P}_B(e^{j\omega}) = \frac{1}{N}\sum_{i=0}^{K-1}\left|\sum_{n=0}^{L-1} x(n+iL)e^{-jn\omega}\right|^2$$

Bias

$$E\left\{\hat{P}_B(e^{j\omega})\right\} = \frac{1}{2\pi}P_x(e^{j\omega}) * W_B(e^{j\omega})$$

Resolution

$$\Delta\omega = 0.89K\frac{2\pi}{N}$$

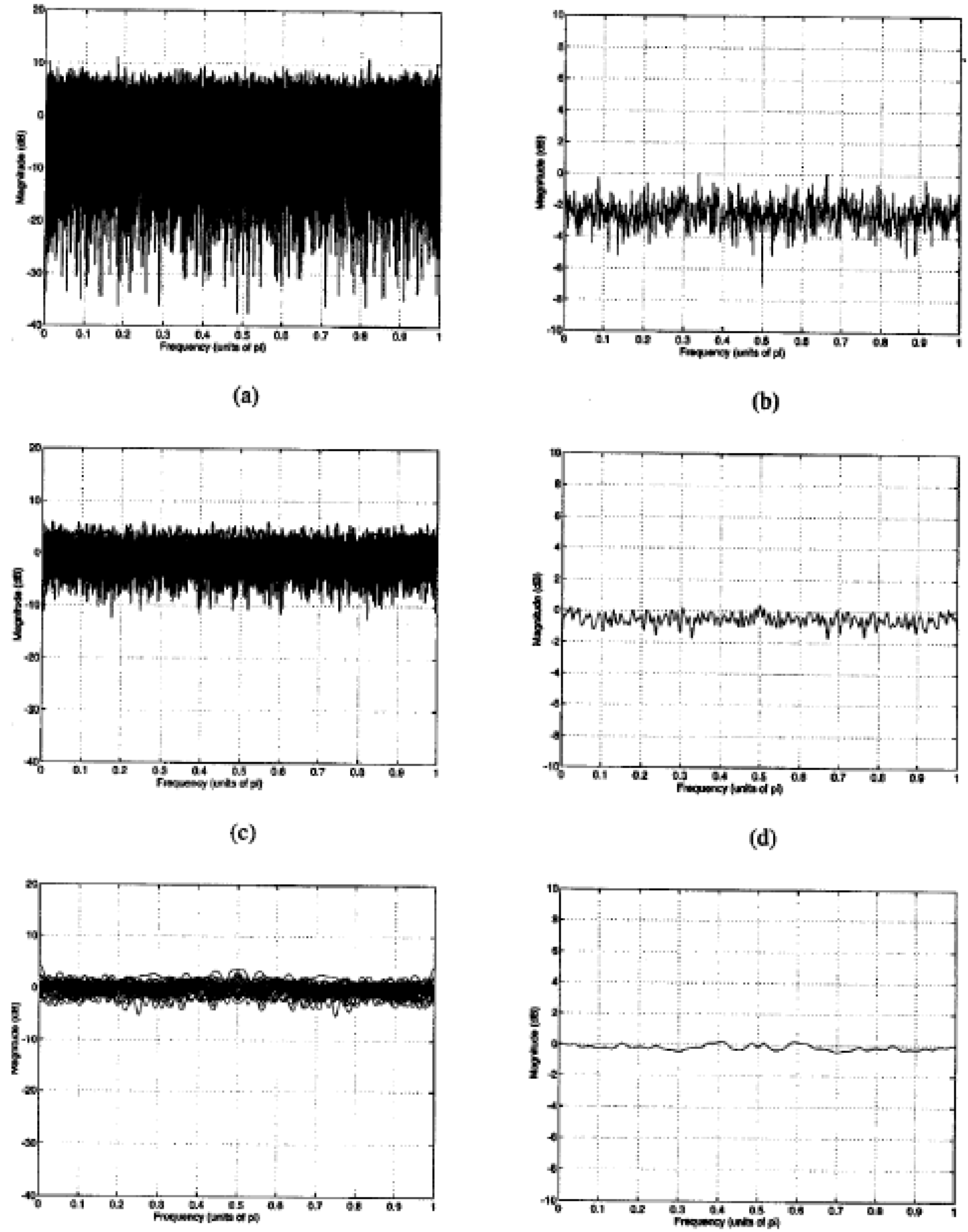
Variance

$$\text{Var}\left\{\hat{P}_B(e^{j\omega})\right\} \approx \frac{1}{K}P_x^2(e^{j\omega})$$

Bartlett's method: White noise

- a) Periodogram with N=512
- b) Ensemble average
- c) Overlay of 50 Bartlett estimates with K=4 and L=128
- d) Ensemble average
- e) Overlay of 50 Bartlett estimates with K=8 and L=64
- f) Ensemble average

Overlay of 50 estimates　　Ensemble Average



(a) (b) (c) (d) (e) (f)

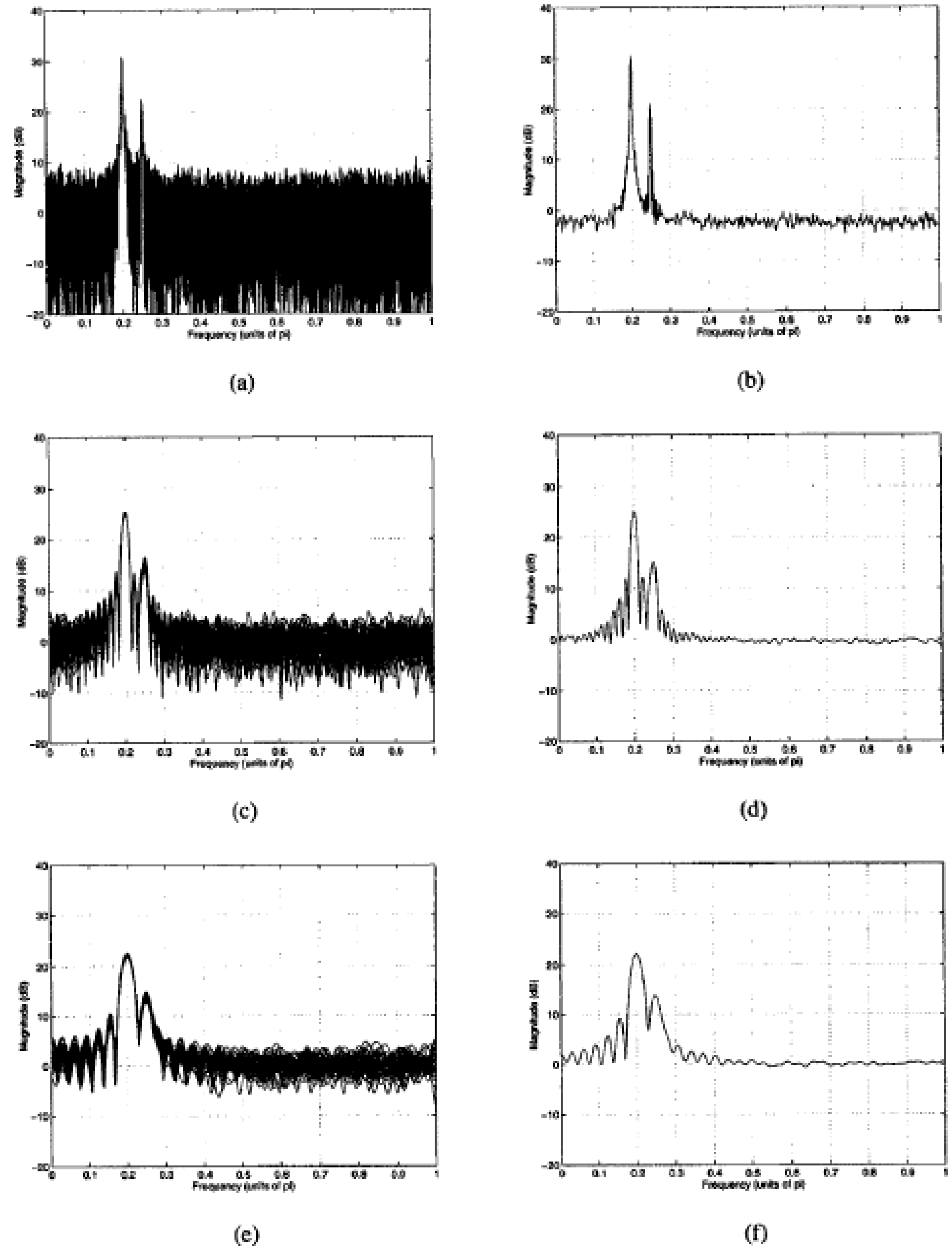
Bartlett’s method: Two sinusoidal in white noise

- a) Periodogram with N=512
- b) Ensemble average

- c) Overlay of 50 Bartlett estimates with K=4 and L=128
- d) Ensemble average

- e) Overlay of 50 Bartlett estimates with K=8 and L=64
- f) Ensemble average

Note how larger K results in shorter L and hence in less spectral resolution



Welch's method

- Two modifications to Bartlett's method
  - 1) the subsequences are allowed to overlap
  - 2) instead of Periodograms, modified Periodograms are averaged

- Assuming that successive sequences are offset by D points and that each sequence is L points long, then the i^{th} sequence is

$$x_i(n) = x(n + iD) \quad ; \quad n = 0, 1, \ldots, L-1$$

- Thus the overlap is L-D points and if K sequences cover the entire N data points then

$$N = L + D(K-1).$$

Welch's method

- For example, with no overlap (D=L) there are K=N/L subsequences of length L
- For a 50% overlap (D=L/2) there is a tradeoff between increasing L or increasing K
  - If L stays the same then there are more subsequences to average, hence the variance of the estimate is reduced

$$K = 2\frac{N}{L} - 1$$

  - If subsequences are doubled in length and hence the spectral resolution is then doubled

$$K = \frac{N}{L} - 1$$

Performance of Welch's method

- Welch's method can be written in terms of the data record as follows

$$\hat{P}_W(e^{j\omega}) = \frac{1}{KLU} \sum_{i=0}^{K-1} \left| \sum_{n=0}^{L-1} w(n)x(n+iD)e^{-jn\omega} \right|^2$$

- Or in terms of modified Periodograms

$$\hat{P}_W(e^{j\omega}) = \frac{1}{K} \sum_{i=0}^{K-1} \hat{P}_M^{(i)}(e^{j\omega})$$

- Hence the expected value of Welch's estimate is

$$E\{\hat{P}_W(e^{j\omega})\} = E\{\hat{P}_M(e^{j\omega})\} = \frac{1}{2\pi LU} P_x(e^{j\omega}) * |W(e^{j\omega})|^2$$

- Where $W(e^{j\omega})$ is the Fourier transform of the L-point data window w(n)

Performance of Welch’s method

- Welch’s method is asymptotically unbiased estimate of the power spectrum

- The variance is much harder to compute, since the overlap results in a correlation

- Nevertheless for an overlap of 50% and a Bartlett window it has been shown that

$$\mathrm{Var}\{\hat{P}_W(e^{j\omega})\} \approx \frac{9}{8K} P_x^2(e^{j\omega})$$

- Recall Bartlett’s Method results in

$$\mathrm{Var}\left\{\hat{P}_B(e^{j\omega})\right\} \approx \frac{1}{K}\mathrm{Var}\left\{\hat{P}_{per}^{(i)}(e^{j\omega})\right\} \approx \frac{1}{K} P_x^2(e^{j\omega})$$

Performance of Welch's method

- For a fixed number of data N, with 50% overlap, twice as many subsequences can be averaged, hence expressing the variance in terms of L and N we have

$$\mathrm{Var}\{\hat{P}_W(e^{j\omega})\} \approx \frac{9}{16}\frac{L}{N}P_x^2(e^{j\omega})$$

- Since N/L is the number of subsequences K used in Bartlett's method it follows

$$\mathrm{Var}\{\hat{P}_W(e^{j\omega})\} \approx \frac{9}{16}\mathrm{Var}\{\hat{P}_B(e^{j\omega})\}$$

- In other words, and not surprising, with 50% overlap (and Bartlett window), the variance of Welch's method is about half that of Bartlett's method

Welch's method summary

Table 8.5 Properties of Welch's Method

$$\hat{P}_W(e^{j\omega}) = \frac{1}{KLU} \sum_{i=0}^{K-1} \left| \sum_{n=0}^{L-1} w(n)x(n+iD)e^{-jn\omega} \right|^2$$

$$U = \frac{1}{L} \sum_{n=0}^{L-1} |w(n)|^2$$

Bias

$$E\left\{\hat{P}_W(e^{j\omega})\right\} = \frac{1}{2\pi LU} P_x(e^{j\omega}) * |W(e^{j\omega})|^2$$

Resolution Window dependent

Variance†

$$\operatorname{Var}\left\{\hat{P}_W(e^{j\omega})\right\} \approx \frac{9}{16}\frac{L}{N} P_x^2(e^{j\omega})$$

† Assuming 50% overlap and a Bartlett window.

An example of Welch's method

Consider the process defined in Example 8.2.5 consisting of two sinusoids in unit variance white noise. Using Welch's method with $N = 512$, a section length $L = 128$, a 50% overlap (7 sections), and a Hamming window, an overlay plot of the spectrum estimates for 50 different realizations of the process are shown in Fig. 8.17*a* and the ensemble average is shown in Fig. 8.17*b*. Comparing these estimates with those shown in Fig. 8.15*e* and *f* of Example 8.2.5, we see that, since the number of sections used in both examples are about the same (7 versus 8), then the variance of the two estimates are approximately the same. In addition, although the width of the main lobe of the Hamming window used in Welch's method is 1.46 times the width of the rectangular window used in Bartlett's method, the resolution is about the same. The reason for this is due to the fact that the 50% overlap that is used in Welch's method allows for the section length to be twice the length of that used in Bartlett's method. So what do we gain with Welch's method? We gain a reduction in the amount of spectral leakage that takes place through the sidelobes of the data window.

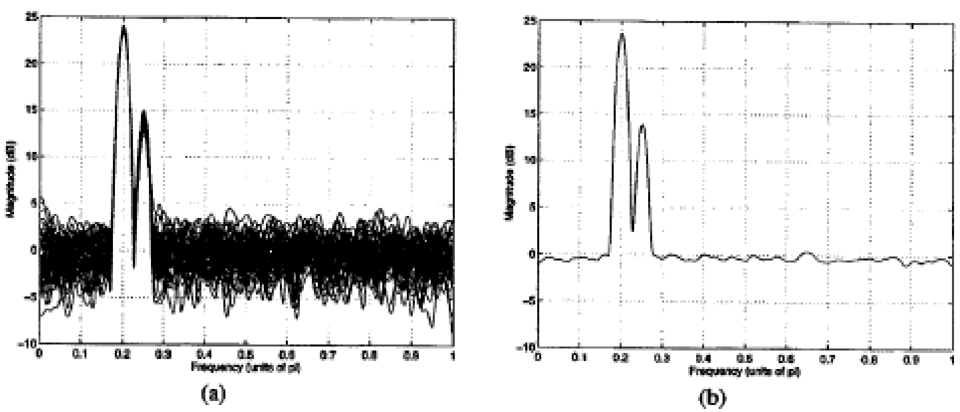


Figure 8.17 *(a) An overlay plot of 50 estimates of the spectrum of two sinusoids in noise using Welch's method with $N = 512$, a section length of $L = 128$, 50% overlap (7 sections), and a Hamming window. (b) The average of the estimates in (a).*

Exercises

Exercise

8.1 Given $N = 10,000$ samples of a process $x(n)$, you are asked to compute the periodogram. However, with only a finite amount of memory resources, you are unable to compute a DFT any longer than 1024. Using these 10,000 samples, describe how you would be able to compute a periodogram that has a resolution of

$$\Delta\omega = 0.89\frac{2\pi}{10000}$$

Hint: Consider how the decimation-in-time FFT algorithm works.

Solution

To get the maximum resolution from $N = 10000$ data values, we want to compute the periodogram of $x(n)$ (segmenting $x(n)$ into subsequences reduces the resolution). The question, therefore, is how to compute the periodogram of $x(n)$ using 1024-point DFT's. Recalling how the FFT works, note that

$$X(e^{j\omega}) = \sum_{n=0}^{9999} x(n)e^{-jn\omega} = \sum_{n=0}^{999}\sum_{l=0}^{9} x(10n+l)e^{-j(10n+l)\omega} = \sum_{l=0}^{9} e^{-jl\omega} \sum_{n=0}^{999} x(10n+l)e^{-jn\omega\cdot 10}$$

Therefore, the procedure is to pad $x(n)$ to form a sequence of length $N = 10240$, and then decimate $x(n)$ into 10 sequences $x_l(n)$ of length $M = 1024$,

$$x_l(n) = x(10n+l) \quad ; \quad n = 0, 1, \ldots, 1023$$

Next, the 1024-point DFT's of these sequences, $X_l(k)$, are computed, and combined using the "twiddle factors" $\exp\left(-jl\frac{2\pi k}{10240}\right)$ as follows

$$X(k) = \sum_{l=0}^{9} e^{-jl\frac{2\pi k}{10240}} X_l(k) \quad ; \quad k = 0, 1, \ldots, 10239$$

Finally, squaring the magnitude of $X(k)$ and dividing by $N = 10240$, we have the periodogram with a resolution $\Delta\omega = 0.89(2\pi/10000)$.

Solution

$N = 1024$

$$X\left(e^{j\frac{2\pi}{10\cdot N}\cdot k}\right) = \sum_{n=0}^{10\cdot N-1} x[n]\cdot e^{-j\frac{2\pi}{10N}k\cdot n}$$

$$= \sum_{m=0}^{N-1}\sum_{l=0}^{9} x[10m+l]\, e^{-j\frac{2\pi}{10N}k\cdot[10m+l]}$$

$$= \sum_{l=0}^{9} e^{-j\frac{2\pi}{10N}k\cdot l}\sum_{m=0}^{N-1} x[10m+l]\, e^{-j\frac{2\pi}{10N}k\cdot 10m}$$

$$= \sum_{l=0}^{9} e^{-j\frac{2\pi}{10N}k\cdot l}\cdot\sum_{m=0}^{N-1} x[10m+l]\cdot e^{-j\frac{2\pi}{N}\cdot k\cdot m}$$

$$x_l^{10}(m) \triangleq x[10m+l]$$

$$= \sum_{l=0}^{9} e^{-j\frac{2\pi}{10N}k l}\underbrace{\sum_{m=0}^{N-1} x_l^{10}(m)\, e^{-j\frac{2\pi}{N}k\cdot m}}_{DFT}$$

$$X\left(e^{j\frac{2\pi}{10N}k}\right) = \sum_{l=0}^{9} e^{-j\frac{2\pi}{10N}k\cdot l}\cdot \underbrace{X_l^{10}[k]}_{}$$

periodic with N

$$X_l^{10}[k] = X_l^{10}[k+N]$$

Exercise

8.2 A continuous-time signal $x_a(t)$ is bandlimited to 5 kHz, i.e., $x_a(t)$ has a spectrum $X_a(f)$ that is zero for $|f| > 5$ kHz. Only 10 seconds of the signal has been recorded and is available for processing. We would like to estimate the power spectrum of $x_a(t)$ using the available data in a radix-2 FFT algorithm, and it is required that the estimate have a resolution of at least 10 Hz. Suppose that we use Bartlett's method of periodogram averaging.

(a) If the data is sampled at the Nyquist rate, what is the minimum section length that you may use to get the desired resolution?

(b) Using the minimum section length determined in part (a), with 10 seconds of data, how many sections are available for averaging?

(c) How does your choice of the sampling rate affect the resolution and variance of your estimate? Are there any benefits to sampling above the Nyquist rate?

Solution

(a) If we sample at the Nyquist rate, $f_s = 10\text{kHz}$, then a resolution of $\Delta f = 10\text{Hz}$ (in analog frequency) implies that we want a resolution (in radians) of

$$\Delta\omega = 2\pi \frac{\Delta f}{f_s} = 2\pi \times 10^{-3}$$

Since the resolution of the periodogram using an L-point data record is

$$\text{Res}[\widehat{P}_{PER}(\omega)] = \Delta\omega = 0.89\frac{2\pi}{L}$$

then for Bartlett's method we want to use a section length of

$$L \geq 0.89\frac{2\pi}{\Delta\omega} = 890 \quad \text{samples}$$

(b) Sampling at 10 kHz, 10 seconds of data corresponds to $N = (10)(10 \times 10^3) = \times 10^5$ samples. Therefore, with a 1024-point DFT the number of sections we may have in Bartlett's method is

$$K = [N/1024] = 98$$

(c) If the sampling rate is increased then $\Delta\omega$ decreases which, in turn, requires a longer section length for a given resolution. However, an increase in the sampling rate produces a corresponding increase in the total number of samples within a T second interval. Therefore, since the variance (normalized) is

$$V = L/N$$

increasing the sampling rate has no effect. Thus, provided that the sampling rate is not less than the Nyquist frequency, the resolution and the variance do not depend on the sampling rate.

Exercise

8.3 Bartlett's method is used to estimate the power spectrum of a process from a sequence of $N = 2000$ samples.

(a) What is the minimum length L that may be used for each sequence if we are to have a resolution of $\Delta f = 0.005$?

(b) Explain why it would not be advantageous to increase L beyond the value found in (a).

(c) The *quality factor* of a spectrum estimate is defined to be the inverse of the variability,

$$Q = 1/\mathcal{V}$$

Using Bartlett's method, what is the minimum number of data samples, N, that are necessary to achieve a resolution of $\Delta f = 0.005$, and a quality factor that is five times larger than that of the periodogram?

Solution

(a) Since $\Delta f = 0.9/L$ then

$$L = \frac{0.89}{\Delta f} = \frac{0.9}{0.005} = 180$$

(b) Increasing L will increase the resolution, but it will also result in a decrease in the number of segments that may be averaged. This, in turn, will increase the variance of the spectrum estimate.

(c) For the periodogram, the quality factor is $Q_{per} = 1/\mathcal{V}_{per} = 1$. The quality factor for Bartlett's method is $Q_B = 1/\mathcal{V}_B = K$. Therefore, if we want $Q_{per}/Q_B \geq 5$, then we must have $K \geq 5$. With $M = 180$ (for $\Delta f = 0.005$), then we must have

$$N = KM \geq 5 \times 180 = 900$$

Exercise

8.5 Many commercial *Fourier analyzers* continuously update the estimate of the power spectrum of a process $x(n)$ by exponential averaging periodograms as follows,

$$\hat{P}_i(e^{j\omega}) = \alpha \hat{P}_{i-1}(e^{j\omega}) + \frac{1-\alpha}{N}\left|\sum_{n=0}^{N-1} x_i(n)e^{-jn\omega}\right|^2$$

where $x_i(n) = x(n+Ni)$ is the ith sequence of N data values. This update equation is initialized with $\hat{P}_{-1}(e^{j\omega}) = 0$.

(a) Qualitatively describe the philosophy behind this method, and discuss how the value for the weighting factor α should be selected.

(b) Assuming that successive periodograms are uncorrelated and that $0 < \alpha < 1$, find the mean and variance of $\hat{P}_i(e^{j\omega})$ for a Gaussian random process.

(c) Repeat the analysis in part (b) if the periodograms are replaced with modified periodograms.

Solution

(a) As data is being read by a spectrum analyzer, the goal is to continuously update the estimate. As each data record of length N is collected, the periodogram is computed, and *averaged* with the previous spectrum estimate. Although a running average could be formed, this would assume that the process is stationary. Selecting a value of $0 < \alpha < 1$ allows the estimate to *forget* $\widehat{P}_i(e^{j\omega})$ as more data is collected. In the extreme case in which $\alpha = 0$, $\widehat{P}_i(e^{j\omega})$ is the periodogram of the most recent N data values. As we will see in part (b), $\widehat{P}_i(e^{j\omega})$ is an exponentially weighted average of the previous periodograms.

Solution

(b) If we define

$$Q_i(e^{j\omega}) = \frac{1}{N}\left|\sum_{n=0}^{N-1} x_i(n)e^{-jn\omega}\right|^2$$

then the expression for the ith spectrum estimate, $P_i(e^{j\omega})$ is

$$\hat{P}_i(e^{j\omega}) = \alpha\hat{P}_{i-1}(e^{j\omega}) + (1-\alpha)Q_i(e^{j\omega})$$

$$y_1 = (1-\alpha)x_1$$
$$y_2 = \alpha \cdot y_1 + (1-\alpha)x_2$$
$$y_3 = \alpha y_2 + (1-\alpha)x_3$$
$$= \alpha^2 y_1 + \alpha(1-\alpha)x_2 + (1-\alpha)x_3$$
$$= \alpha^2(1-\alpha)x_1 + \alpha(1-\alpha)x_2 + (1-\alpha)x_3$$
$$\Rightarrow y_i = \sum_{k=0}^{i}(1-\alpha)\cdot\alpha^k\cdot x_{i-k}$$

which is a difference equation for $\hat{P}_i(e^{j\omega})$. Since the initial conditions are zero, $\hat{P}_{-1}(e^{j\omega}) = 0$, then the solution for $\hat{P}_i(e^{j\omega})$ is

$$\hat{P}_i(e^{j\omega}) = \sum_{k=0}^{i}(1-\alpha)\alpha^k Q_{i-k}(e^{j\omega})$$

Taking the expected value we have

$$E\{\hat{P}_i(e^{j\omega})\} = \sum_{k=0}^{i}(1-\alpha)\alpha^k E\{Q_{i-k}(e^{j\omega})\}$$

Since $Q_{i-k}(e^{j\omega})$ is the periodogram of $x_{i-k}(n)$, then

$$E\{Q_{i-k}(e^{j\omega})\} = \frac{1}{2\pi}P_x(e^{j\omega}) * W_B(e^{j\omega})$$

← independent of $(i-k)$

Solution

and

$$
\begin{aligned}
E\{\widehat{P}_i(e^{j\omega})\} &= \frac{1}{2\pi}\sum_{k=0}^{i}(1-\alpha)\alpha^k\left[P_x(e^{j\omega}) * W_B(e^{j\omega})\right] \\
&= \frac{1}{2\pi}\left[P_x(e^{j\omega}) * W_B(e^{j\omega})\right](1-\alpha)\sum_{k=0}^{i}\alpha^k \\
&= \frac{1}{2\pi}\left[P_x(e^{j\omega}) * W_B(e^{j\omega})\right](1-\alpha)\frac{1-\alpha^{i+1}}{1-\alpha} \\
&= (1-\alpha^{i+1})\frac{1}{2\pi}\left[P_x(e^{j\omega}) * W_B(e^{j\omega})\right]
\end{aligned}
$$

$$\sum_{n=0}^{N-1} q^n = \frac{1-q^N}{1-q} \qquad |q|<1$$

For the variance, we proceed in the same way, using the fact that the variance of the periodogram is

$$\text{var}\{\widehat{P}_{\text{per}}(e^{j\omega})\} \approx P_x^2(e^{j\omega})$$

Therefore, we have

$$
\begin{aligned}
\text{var}\{\widehat{P}_i(e^{j\omega})\} = \sout{E\{\widehat{P}_i(e^{j\omega})\}} &= \sum_{k=0}^{i}(1-\alpha)^2\alpha^{2k}P_x^2(e^{j\omega}) = (1-\alpha)^2 P_x^2(e^{j\omega})\sum_{k=0}^{i}\alpha^{2k} \\
&= (1-\alpha)^2\left[\frac{1-\alpha^{2(i+1)}}{1-\alpha^2}\right]P_x^2(e^{j\omega}) = \frac{1-\alpha}{1+\alpha}(1-\alpha^{2(i+1)})P_x^2(e^{j\omega})
\end{aligned}
$$

Solution

(c) For modified periodograms, the only change that is necessary is to use

$$E\{Q_i(e^{j\omega})\} = \frac{1}{NU}\left|P_x(e^{j\omega}) * W_B(e^{j\omega})\right|^2$$

where

$$U = \frac{1}{N}\sum_{n=0}^{N-1}|w(n)|^2$$

Substituting this into the expression in part (b), we have for the expected value,

$$E\{\widehat{P}_i(c^{j\omega})\} = (1-\alpha^{i+1})\frac{1}{2\pi NU}P_x(e^{j\omega}) * |W_B(e^{j\omega})|^2$$

and the variance is the same.