$$\boxed{\text{Capon's Method}}$$

Sinusoid Parameter Estimation

For the moment let's forget about spectrum estimation. Consider the complex sinusoid

$$x[n] = A e^{2\pi i f_0 n + i\phi} + z[n], \qquad n = 0, 1, \ldots, N-1$$

where

- $\{z[n]\}$ is complex Gaussian with covariance Γ_{zz}
- f_0 is known
- A, ϕ are unknown and desired

Maximum Likelihood Estimation

Suppose $\underline{Y}$ is a random variable/vector with pdf $f(\underline{y} \mid \underline{\theta})$. Suppose $\underline{y}$ is a realization of $\underline{Y}$. The *likelihood* function for $\underline{\theta}$ is

$$\ell(\underline{\theta} / \underline{y}) = f(\underline{y} / \underline{\theta}).$$

The *maximum likelihood estimate* of $\underline{\theta}$ is

$$\hat{\underline{\theta}} := \arg\max_{\underline{\theta}} \ell(\underline{\theta} / \underline{y})$$

The MLE has many nice properties such as asymptotic efficiency (unbiased, variance = cramer-rao lower bound).

Write $A_c = Ae^{j\phi}$, so that

$$x[n] = A_c e^{2\pi i f_0 n} + z[n].$$

In vector notation,

$$\underline{x} = A_c \underline{e} + \underline{z}$$

where

$$\underline{x} = [x[0] \quad \cdots \quad x[N-1]]^T$$

$$\underline{e} = [1 \quad e^{2\pi i f_0} \quad \cdots \quad e^{2\pi i f_0 (N-1)}]^T$$

$$\underline{z} = [z[0] \quad \cdots \quad z[N-1]]^T$$

The density of $\underline{z}$ is

$$f(\underline{z}) = \frac{1}{\pi^N |\Gamma_{zz}|} \exp\left\{ -\underline{z}^H \Gamma_{zz}^{-1} \underline{z} \right\}$$

What is the likelihood function for A_c?

$$\ell(A_c \mid \underline{x}) = \frac{1}{\pi^N |\Gamma_{zz}|} \exp\left\{ -(\underline{x} - A_c \underline{e})^H \Gamma_{zz}^{-1} (\underline{x} - A_c \underline{e}) \right\}$$

To maximize the likelihood w.r.t A_c we need to minimize

$$(\underline{x} - A_c \underline{e})^H \Gamma_{zz}^{-1} (\underline{x} - A_c \underline{e})$$

w.r.t. A_c.

Some enjoyable algebra reveals

$$(\underline{x} - A_c \underline{e})^H \Gamma_{zz}^{-1} (\underline{x} - A_c \underline{e}) =$$

$$(\underline{x} - \tilde{A}_c \underline{e})^H \Gamma_{zz}^{-1} (\underline{x} - \tilde{A}_c \underline{e}) \quad +$$

$$|A_c - \tilde{A}_c|^2 \, \underline{e}^H \Gamma_{zz}^{-1} \underline{e}$$

where

$$\tilde{A}_c = \frac{\underline{e}^H \Gamma_{zz}^{-1} \underline{x}}{\underline{e}^H \Gamma_{zz}^{-1} \underline{e}} .$$

So what is the MLE of A_c?

$$\hat{A}_c = \frac{\underline{e}^H \Gamma_{zz}^{-1} \underline{x}}{\underline{e}^H \Gamma_{zz}^{-1} \underline{e}}$$

<u>Observe</u>: $\hat{A}_c$ is a <u>linear</u> estimator

$$\hat{A}_c = \underline{w}^H \underline{x}$$

where $$\underline{w} = \frac{\Gamma_{zz}^{-1} \underline{e}}{\underline{e}^H \Gamma_{zz}^{-1} \underline{e}}.$$

LMVU Estimation

A second method for estimating A_c is called <u>linear</u> <u>minimum</u> <u>variance</u> <u>unbiased</u> estimation:

$$\hat{A}_c = \hat{\underline{w}}^H \underline{x}$$

where

$$\hat{\underline{w}} = \underset{\underline{w}}{\operatorname{argmin}} \; Var\{\underline{w}^H \underline{x}\}$$

$$\text{s.t.} \quad E\{\underline{w}^H \underline{x}\} = A_c .$$

<u>Exercise</u> Evaluate $E\{\underline{w}^H \underline{x}\}$ and find a necessary and sufficient condition on $\underline{w}$ for $\underline{w}^H \underline{x}$ to be unbiased.

Solution

$$E\{\underline{w}^H \underline{x}\} = E\{\underline{w}^H (A_c \underline{e} + \underline{z})\}$$

$$= \underline{w}^H \underline{e} A_c + \underline{w}^H \cancelto{0}{E\{\underline{z}\}}$$

$$= \underline{w}^H \underline{e} A_c$$

Therefore, for $\underline{w}^H \underline{x}$ to be unbiased, we need

$$\boxed{\underline{w}^H \underline{e} = 1}$$

What about the variance?

$$\mathrm{Var}\{\underline{w}^H x\} = E\{(\underline{w}^H \underline{x} - \underbrace{\underline{w}^H \underline{e}}_{= 1} A_c)^2\}$$

$$= E\{\underline{w}^H \underbrace{(\underline{x} - A_c \underline{e})}_{\underline{z}} \cdot \underbrace{(x - A_c \underline{e})^H}_{\underline{z}^H} \underline{w}\}$$

$$\boxed{= \underline{w}^H \Gamma_{zz} \underline{w}}$$

Therefore, the LMVUE is

$$\hat{A}_c = \hat{\underline{w}}^H \underline{x}$$

where

$$\boxed{\begin{aligned} \hat{\underline{w}} = \arg\min_{\underline{w}} \quad & \underline{w}^H \Gamma_{zz} \underline{w} \\ \text{s.t.} \quad & \underline{w}^H \underline{e} = 1 \end{aligned}}$$

We can solve this constrained optimization problem using <u>Lagrange</u> <u>multipliers</u>.

Define the <u>Lagrangian</u>

$$L(\underline{w}, \lambda) = \underline{w}^H \Gamma_{zz} \underline{w} + \lambda (\underline{w}^H \underline{e} - 1)$$

Define

$$\hat{\underline{w}}(\lambda) = \arg\min_{\underline{w}} L(\underline{w}, \lambda)$$

Proposition: If λ^* is such that

$$\underline{\hat{w}}(\lambda^*)^H \underline{e} = 1$$

then $\underline{\hat{w}}(\lambda^*) = \underline{\hat{w}}$ (the LMVUE)

Let's apply this result.

$$\frac{\partial L}{\partial \underline{w}} = 2\Gamma_{zz}\underline{w} + \lambda \underline{e} = 0$$

$$\Rightarrow \boxed{\underline{\hat{w}}(\lambda) = -\frac{\lambda}{2}\Gamma_{zz}^{-1}\underline{e}}$$

To find λ^* we set

$$\underline{\hat{w}}(\lambda)^H \underline{e} = 1$$

$$\Rightarrow -\frac{\lambda}{2}\underline{e}^H \Gamma_{zz}^{-1}\underline{e} = 1$$

$$\Rightarrow \lambda^* = \frac{-2}{\underline{e}^H \Gamma_{zz}^{-1}\underline{e}}$$

$$\Rightarrow \boxed{\underline{\hat{w}} = \frac{\Gamma_{zz}^{-1}\underline{e}}{\underline{e}^H \Gamma_{zz}^{-1}\underline{e}}}$$

Conclusion: The MLE and LMVUE are the same.

Exercise: What is the variance of $\hat{A}_c$?

Solution:

$$\mathrm{Var}\{\hat{A}_c\} = \underline{\hat{w}}^H \Gamma_{zz} \underline{\hat{w}}$$

$$= \frac{\underline{e}^H \Gamma_{zz}^{-1}}{\underline{e}^H \Gamma_{zz}^{-1} \underline{e}} \cdot \Gamma_{zz} \cdot \frac{\Gamma_{zz}^{-1} \underline{e}}{\underline{e}^H \Gamma_{zz}^{-1} \underline{e}}$$

$$= \boxed{\frac{1}{\underline{e}^H \Gamma_{zz}^{-1} \underline{e}}}$$

Filtering Interpretation of MLE / LMVUE

$$x[n] \longrightarrow \boxed{W(z)} \longrightarrow y[n]$$

We may view the LMVUE as an FIR filter:

$$\hat{A}_c = y[N-1] = \hat{\underline{w}}^H \cdot \big[x[0] \; \cdots \; x[N-1] \big]^T$$

FIR filter

The unbiasedness of $\hat{\underline{w}}$ means

$$\begin{aligned} 1 &= \underline{w}^H \underline{e} \\ &= \sum_{n=0}^{N-1} w^*[n]\, e^{-2\pi i f_0 n} \\ &= W(f_0) \end{aligned}$$

That is, the filter passes the sinusoid at frequency f_0 <u>without</u> <u>distortion</u>.

The minimum variance property means that the power of the output signal is minimized. (subject to the unbiased constraint).

But the noise power of the output is

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} P_{xx}(f)\,|W(f)|^2\,df$$

$\Rightarrow$ $W(f)$ tries to minimize the frequency response at $f \neq f_0$.

<u>Example</u> $z[n]$ is white Gaussian noise, variance σ_z^2.

Then

$$\hat{\underline{w}} = \frac{\Gamma_{zz}^{-1}\,\underline{e}}{\underline{e}^H\,\Gamma_{zz}^{-1}\,\underline{e}} =$$

So

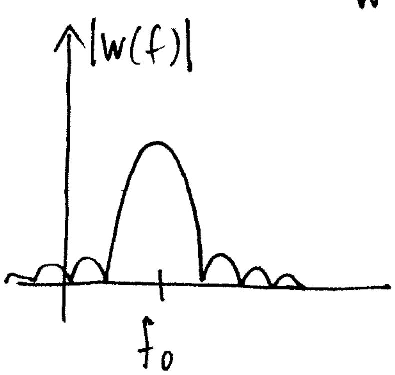
$$\hat{A}_c = \hat{\underline{w}}^H\,\underline{x}$$

$$= \frac{1}{N}\,\underline{e}^H\,\underline{x}$$

$$= \frac{1}{N}\sum_{n=0}^{N-1} x[n]\,e^{-2\pi i f_0 n}$$

$$= \text{DTFT of } \underline{x} \text{ at } f_0$$

From the fitering perspective



$$W(f) = \frac{1}{N}\sum_{n=0}^{N-1} e[n]\,e^{-2\pi i f n}$$

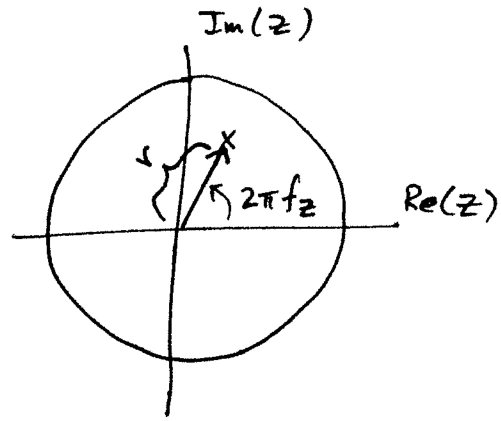
$$= \frac{1}{N}\sum_{n=0}^{N} e^{-2\pi i (f-f_0) n}$$

$$= \frac{\sin[N\pi(f-f_0)]}{N\sin[\pi(f-f_0)]}\cdot \exp\{i(N-1)\pi(f-f_0)\}$$

= "digital sinc centered at f_0"

Example 2 $z[n]$ is AR(1) with

$$-a[1] = re^{2\pi i f_z}$$

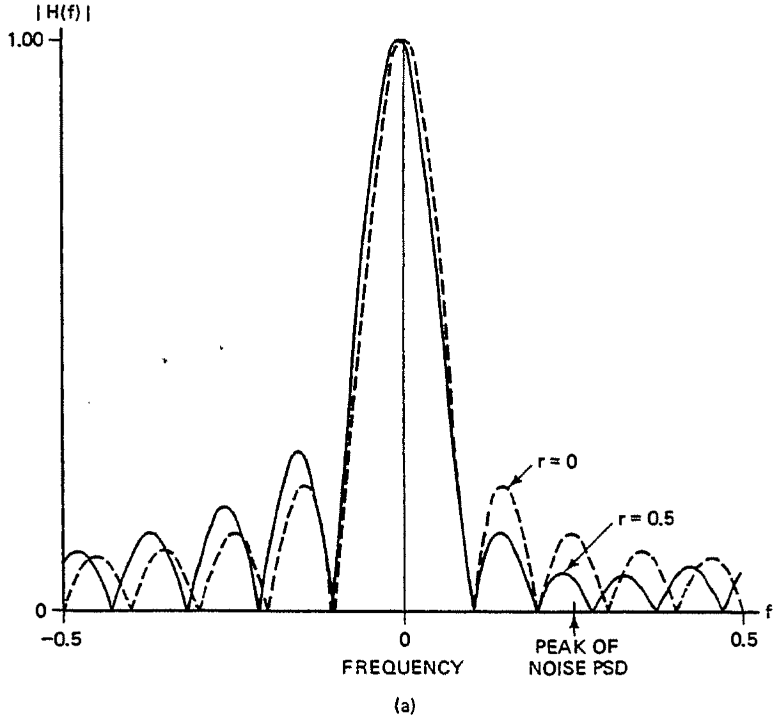


The covariance of the noise is now

$$\Gamma_{zz}^{-1} = \begin{bmatrix} 1 & a^*[1] & 0 & \cdots & 0 \\ a[1] & 1+|a[1]|^2 & a^*[1] & \cdots & 0 \\ 0 & a[1] & 1+|a[1]|^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & a^*[1] \\ 0 & 0 & \cdots & a[1] & 1 \end{bmatrix}$$

(smaller $r \implies$ closer to white noise)

The result is that $W(f)$ now makes an extra effort to suppress the frequency content of $\underline{x}$ near f_z.



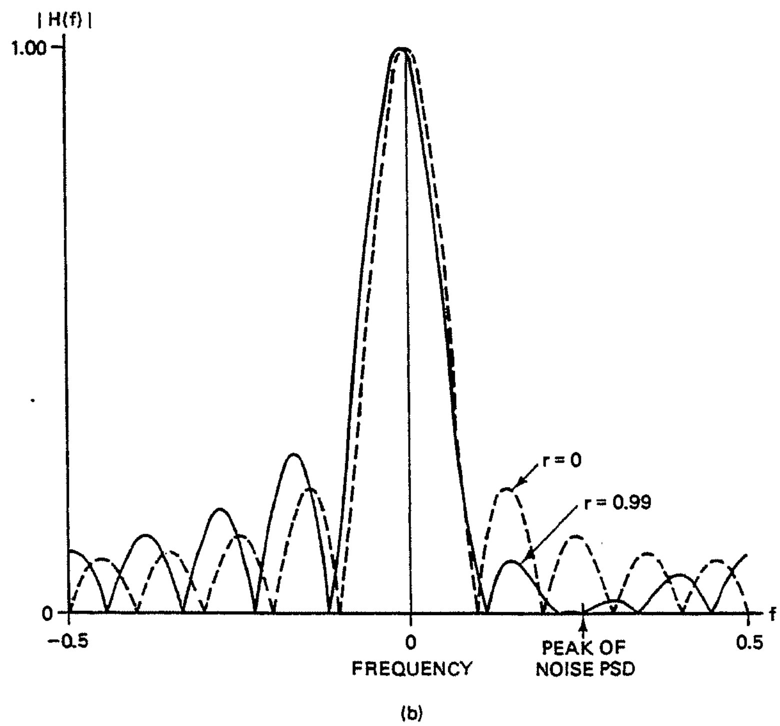


Figure 11.2 Frequency response magnitude of LMVU filter.

Conclusions: The LMVUE filter $W(f)$ passes frequency content at f_0 without distortion while optimally adapting to suppress the frequency content of the noise.

$$\boxed{\text{Capon's Method}}$$

Also called the "minimum variance spectral estimator" (MVSE)

Recall the filtered signal variance

$$\mathrm{Var}\{\hat{A}_c\} = \frac{1}{\underline{e}^H \Gamma_{zz}^{-1} \underline{e}}$$

estimates the power of the signal in the frequency band around f_0.

Now, make 2 modifications:

1) Assume $x[n]$ is a generall WSS RP.

$$\Gamma_{zz} \longrightarrow \Gamma_{xx}$$

Filtering with $W(f)$ still selects frequency content near f_0, and therefore

$$\frac{1}{\underline{e}^H \Gamma_{xx}^{-1} \underline{e}}$$

is still an estimate of the spectrum at f_0.

2) Repeat at all $f_0 \in [-\frac{1}{2}, \frac{1}{2}]$ to get

$$\hat{P}_{CAPON}(f) = \frac{1}{\underline{e}(f)^H \hat{R}_{xx}^{-1} \underline{e}(f)}$$

where

$$\underline{e}(f) = \left[1 \quad e^{2\pi i f} \quad \cdots \quad e^{2\pi i f \cdot p}\right]^T$$

Note p is a parameter that needs to be set. $\hat{R}_{xx}$ is a $p \times p$ ACV matrix determined by one of the usual methods.

Normalization: To ensure proper normalization/scaling of the estimate, use

$$\hat{P}_{CAPON}(f) = \frac{P}{\underline{e}(f)^H \hat{R}_{xx}^{-1} \underline{e}(f)}$$

Properties:

- For estimating PSD of superposition of sinusoids, CAPON has less resolution than AR but more then periodogram methods.

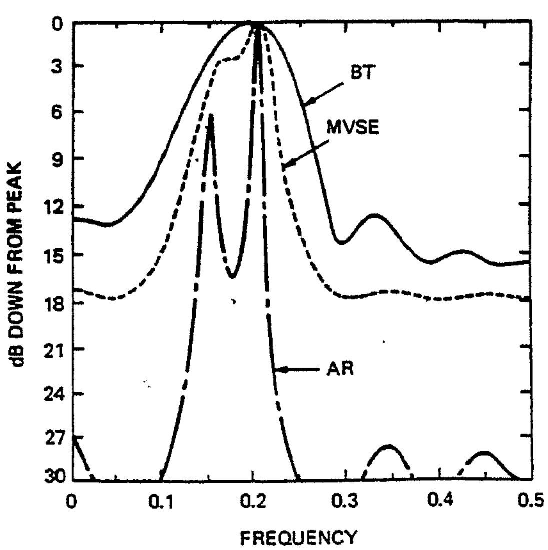


Figure 11.3 Comparison of various spectral estimators. (After Lacoss [1971].)

- Neat fact:

$$\frac{1}{\hat{P}_{CAPON}(f)} = \sum_{i=0}^{P} \frac{1}{\hat{P}_{AR}^{(i)}(f)}$$