Subspace Analysis Methods for Line Spectra

In applications such as communications, radar, sonar, and geophysics, the signals of interest are often well described by the following <u>sinusoidal model</u>:

$$x(n) = \sum_{i=1}^{p} \alpha_i e^{j(2\pi f_i n + \phi_i)}$$

where $x(n)$ is a noise-free, complex-valued sinusoidal process with unknown amplitudes $\{\alpha_i\}$ and frequencies $\{f_i\}$ and statistically independent, random phases $\{\phi_i\}$, uniformly distributed on $[0, 2\pi)$. We measure/observe a noisy version of x:

$$y(n) = x(n) + w(n)$$

additive observation noise

The noise $\{w(n)\}$ is assumed to be (complex-valued) circular white noise.

The covariance function and power spectral density of the noisy signal y can be calculated as follows.

Because the phases are iid, we have

$$E\left[e^{j\phi_k} e^{-j\phi_\ell}\right] = 1, \quad k=\ell$$

and

$$E\left[e^{j\phi_k} e^{-j\phi_\ell}\right] = E\left[e^{j\phi_k}\right] \cdot E\left[e^{-j\phi_\ell}\right]$$

$$= \int_0^{2\pi} \frac{e^{j\phi_k}}{2\pi} d\phi_k \int_0^{2\pi} \frac{e^{-j\phi_\ell}}{2\pi} d\phi_\ell$$

$$= 0, \quad k \neq \ell$$

Thus,

$$E\left[e^{j\phi_k} e^{-j\phi_\ell}\right] = \delta(k-\ell)$$

Let $x_k(n) = \alpha_k e^{j(2\pi f_k n - \phi_k)}$. Then

$$E\left[x_k(n)\, x_\ell^*(n-m) \right] = \alpha_k^2 e^{j 2\pi f_k m} \delta(k-\ell)$$

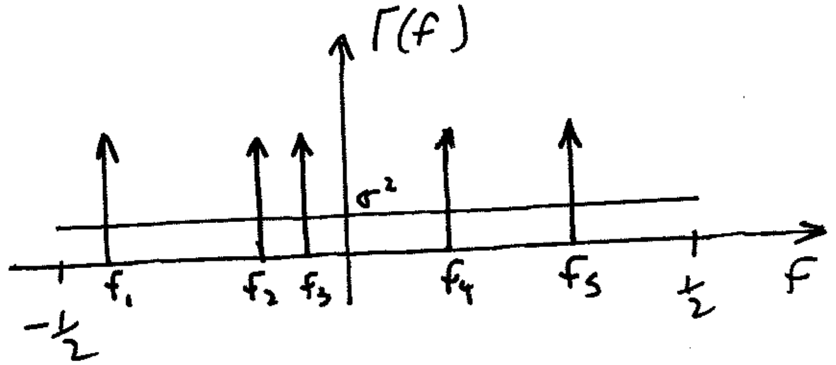
and

$$E\left[y(n)\, y^*(n-m) \right] = \sum_{i=1}^{p} \alpha_i^2 e^{j 2\pi f_i m} + \sigma^2 \delta(m)$$

$$\Rightarrow \boxed{\gamma(m) = \sum_{i=1}^{p} \alpha_i^2 e^{j 2\pi f_i m} + \sigma^2 \delta(m)}$$

The power spectral density is the DTFT of γ:

$$\boxed{\Gamma(f) = \sum_{i=1}^{p} \alpha_i^2 \delta(f-f_i) + \sigma^2}$$



Because the spectrum consists of Dirac impulses, $\Gamma(f)$ is called a <u>line spectrum</u>.

Notice that spectral analysis, in this case, reduces to a parameter estimation problem:

$$\boxed{\text{From observations } y(n),\ n = \ ,\ldots, N-1 \text{ estimate } \{\alpha_i, f_i\} \text{ (and possibly } p\text{).}}$$

Notice also that if p and $\{f_i\}$ are known, then estimation of $\{\alpha_i\}$ is a simple <u>linear regression</u> problem:

$$\begin{bmatrix} y(0) \\ y(1) \\ \cdot \\ \cdot \\ y(N-1) \end{bmatrix} = \begin{bmatrix} 1 & \cdots & 1 \\ e^{j2\pi f_1} & & e^{j2\pi f_p} \\ \vdots & & \vdots \\ e^{j2\pi f_1(N-1)} & & e^{j2\pi f_p(N-1)} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \cdot \\ \cdot \\ \beta_p \end{bmatrix} + \begin{bmatrix} w(0) \\ w(1) \\ \cdot \\ \cdot \\ w(N-1) \end{bmatrix}$$

where $\{\beta_i \equiv \alpha_i e^{j\phi_i}\}$. $\underline{\hat{\beta}}$ given by LS solution.

1) ★

Jointly estimating $\{\beta_i, f_i\}$ is a highly nonlinear regression problem. An intuitively appealing approach is to estimate $\{\beta_i, f_i\}$ as the minimizers of

$$C(\{\beta_i, f_i\}) = \sum_{n=0}^{N-1} \left| y(n) - \sum_{i=1}^{p} \beta_i e^{j2\pi f_i n} \right|^2 .$$

This criterion is called nonlinear least squares. Unfortunately, the minimization is very complicated and requires sophisticated numerical methods. Moreover, $C(\{\beta_i, f_i\})$ has a multimodal surface and there is no guarantee that numerical methods will produce the least squares result (

So, rather than considering this joint optimization, we will focus on the problem of frequency estimation. That is, we'll first estimate $\{f_i\}$, then plug them into $*$ and solve the linear equations for $\{\beta_i\}$.

Covariance Matrix Model

Let $M < N$

$$\underline{s}_i = \left[1 \quad e^{-j2\pi f_i} \quad \cdots \quad e^{-j2\pi f_i (M-1)} \right]^T$$

$$i = 1, \ldots, P \leq M$$

and

$$\underline{y} = \left[y(n) \quad y(n-1) \quad \cdots \quad y(n-M+1) \right]^T.$$

$$\underline{y} = \left[\underline{s}_1 \cdots \underline{s}_P \right] \underline{\beta} + \underline{w}$$

$$= \sum_{i=1}^{P} \beta_i \underline{s}_i + \underline{w}$$

The covariance matrix of $\underline{y}$ is

$$\underline{R} = E\left[\underline{y}\, \underline{y}^* \right]$$

$$= \sum_{k=1}^{P} \sum_{l=1}^{P} E\left[\beta_k \beta_l^* \right] \underline{s}_k \underline{s}_l^H + E\left[\underline{w}\, \underline{w}^* \right]$$

$$= \sum_{k=1}^{P} |\alpha_k|^2 \underline{s}_k \underline{s}_k^H + \sigma^2 \underline{I}$$

Let $\underline{S} = [\underline{s}_1 \dots \underline{s}_p]$, then

$$\boxed{\underline{R} = \underline{S}\,\underline{P}\,\underline{S}^H + \sigma^2 \underline{I}}$$

Where $\underline{P} = \begin{bmatrix} |\alpha_1|^2 & & 0 \\ & \ddots & \\ 0 & & |\alpha_p|^2 \end{bmatrix}$

Note that

$\underline{S}\,\underline{P}\,\underline{S}^H$ is $M \times M$

but its rank is $p \leq M$.

If $p < M$, than smallest $M-p$ eigenvalues of $\underline{R}$ are equal to σ^2. As we'll see shortly, the <u>eigenstructure</u> of $\underline{R}$ contains complete information about the frequencies $\{f_i\}$.

Pisarenko and MUSIC methods

The MUtiple SIgnal Characterization (MUSIC) method (Schmidt 1979, Bienvenu 1979) and Pisarenko's method (Pisarenko 1973, a special case of MUSIC) are derived from the covariance model with $p < N$.

Let $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_M$ denote the <u>eigenvalues</u> of $\underline{R}$ and let $\{\underline{v}_1, \ldots, \underline{v}_M\}$ denote the corresponding <u>orthonormal</u> <u>eigenvectors</u>.

Since $\underline{S}\,\underline{P}\,\underline{S}^H$ has rank p, it follows that it has p strictly positive eigenvalues (the remaining $M-p$ are zero). Let $\{\tilde{\lambda}_i\}_{i=1}^{p}$ denote the positive eigenvalues of $\underline{S}\,\underline{P}\,\underline{S}^H$.

Let $\{\underline{u}_1, \ldots, \underline{u}_p\}$ be the eigenvectors associated with the p non-zero eigenvalues of $\underline{S}\,\underline{P}\,\underline{S}^H$. Consider

$$\underline{u}_i^H \underline{R}\, \underline{u}_i = \underline{u}_i^H \underline{S}\,\underline{P}\,\underline{S}^H \underline{u}_i + \sigma^2 \underline{u}_i^H \underline{I}\, \underline{u}_i$$

$$= \tilde{\lambda}_i + \sigma^2$$

Also,

$$\underline{R}\, \underline{u}_i = \tilde{\lambda}_i \underline{u}_i + \sigma^2 \underline{u}_i$$

$$= \left(\tilde{\lambda}_i + \sigma^2\right) \underline{u}_i \quad .$$

Thus, we see that $\{\underline{u}_i\}_{i=1}^{p}$ are orthonormal eigenvectors of $\underline{R}$ with eigenvalues $\lambda_i = \tilde{\lambda}_i + \sigma^2$.

Let us make the identification

$$\{\underline{v}_1, \ldots \underline{v}_p\} \equiv \{\underline{u}_1, \ldots, \underline{u}_p\}.$$

Let $\underline{G} = [\, \underline{v}_{M-p} , \ldots , \underline{v}_M \,]$.

Then, from #, we have

$$\underline{S}^H \underline{G} = 0$$

Recall that

$$\underline{S} = \underline{S}\left(\{f_i\}_{i=1}^{p}\right)$$

$$= \begin{bmatrix} 1 & & 1 \\ e^{j2\pi f_1} & \cdots & e^{j2\pi f_p} \\ \vdots & & \vdots \\ e^{j2\pi f_1(M-1)} & & e^{j2\pi f_p(M-1)} \end{bmatrix}$$

$$= [\, \underline{S}(f_1) \ \cdots \ \underline{S}(f_2) \,]$$

where $\underline{S}(f_i) = \begin{bmatrix} 1 \\ e^{j2\pi f_i} \\ \vdots \\ e^{j2\pi f_i(M-1)} \end{bmatrix}$

Now, since all of $\underline{R}$'s eigenvectors are orthonormal, we have

$$\underline{v}_k^H \underline{v}_\ell = \delta(k-\ell)$$

and in particular for $k=1,\dots,p$ and $\ell = M-p,\dots,M$

$$v_k^H v_\ell = u_k^H v_\ell = 0.$$

Therefore

$$\underline{v}_i^H \underline{S}\,\underline{P}\,\underline{S}^H \underline{v}_i = 0, \quad i = M-p,\dots,M \qquad \#$$

and

$$\underline{R}\,\underline{v}_i = \sigma^2 \underline{v}_i, \quad i = M-p,\dots,M$$

$$\boxed{\begin{aligned} \mathrm{Span}\left(\{\underline{v}_i\}_{i=1}^{p}\right) &\equiv \text{"signal+noise" subspace} \\ \mathrm{Span}\left(\{v_i\}_{i=M-p}^{M}\right) &\equiv \text{"noise-only" subspace} \end{aligned}}$$

From $\underline{S}^H \underline{G} = 0$ we obtain the following key result.

$\boxed{\text{The true frequency values } \{f_i\}_{i=1}^{p} \text{ are the only solutions of the equation } \underline{s}^H(f)\,\underline{G}\,\underline{G}^H \underline{s}(f) = 0}$

Note: we require that $M > p$ (i.e., we have a noise-only subspace)

Using this result, we can derive the MUSIC algorithm.

MUSIC

Data $y(n)$, $n = 1, \ldots, N$

Step 1: Compute the sample covariance matrix

$$\underline{\hat{R}} = \frac{1}{N} \sum_{n=M}^{N} \underline{y}_n \underline{y}_n^H$$

where $\underline{y}_n = [y(n), y(n-1), \ldots, y(n-M+1)]^T$

Step 2a: (Spectral Music) (Schmidt '79)

Determine frequency estimates as the locations of the p highest peaks of

$$\frac{1}{\underline{s}^H(f)\, \underline{\hat{G}}\, \underline{\hat{G}}^H \underline{s}(f)} \qquad f \in [-\tfrac{1}{2}, \tfrac{1}{2}]$$

$\underline{\hat{G}}$ defined similarly to $\underline{G}$, but based on eigenvecs of $\underline{\hat{R}}$

- OR -

Step 2b: (Root Music) (Barabell '83)

Determine frequency estimates as the angular positions of the p (pairs of reciprocal) roots of

$$\underline{s}(z^{-1})\, \underline{\hat{G}}\, \underline{\hat{G}}^H \underline{s}(z^{-1}) = 0$$

which are closest to the unit circle.

$$s(z) = [1 \;\; z^{-1} \; \cdots \; z^{-(M-1)}]^T$$

MUSIC requires a noise-only subspace in order to work. Therefore, the minimum value for M is $M = p+1$ (1-dim noise-only subspace).

For $M = p+1$, the MUSIC algorithm reduces to the Pisarenko method. In this case, we use the Step 2b approach which boils down to the degree p polynomial equation

$$\underline{s}^T(z^{-1})\,\hat{\underline{v}}_M = 0$$

where $\hat{\underline{v}}_M$ is the eigenvector associated with the smallest eigenvalue of $\hat{\underline{R}}$.

Ex. $$y(n) = \sum_{i=1}^{4} \alpha_i e^{j(2\pi f_i n + \phi_i)} + w(n)$$

$\alpha_i = 1, \quad i = 1, 2, 3, 4$

$\{w(n)\}$ zero-mean white noise

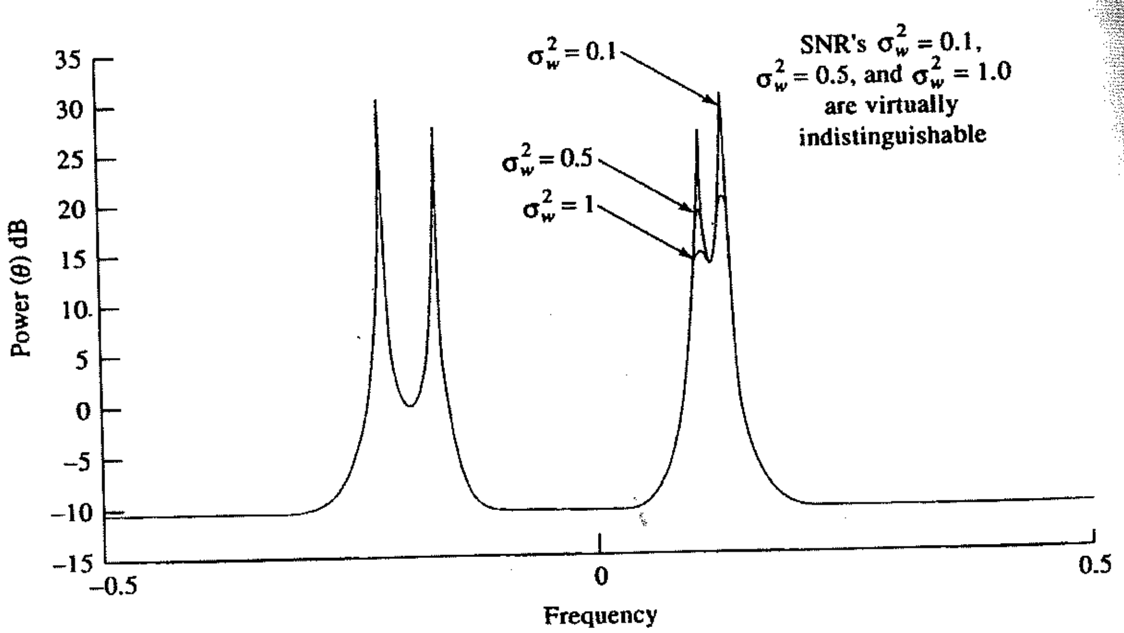
$\{\phi_i\}$ iid uniform $(0, 2\pi)$

$f_1 = -0.222, \ f_2 = -0.166, \ f_3 = 0.10, \ f_4 = 0.1222$

$N = 1024$

MUSIC "Pseudospectrum"

$$\Gamma_{music}(f) = \frac{1}{\sum_{i=p+1}^{M} |\underline{s}^H(f)\, \underline{v}_i|^2}$$



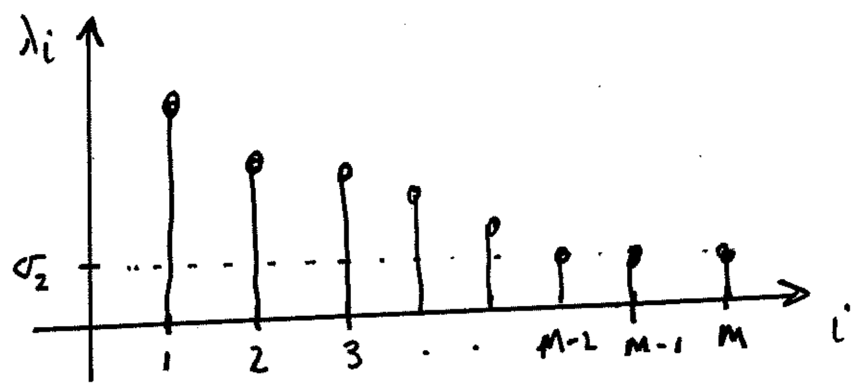
Model Order Selection

Recall that if the signal y consists of p sinusoids in white noise, then the covariance matrix has eigenvalues

$$\{\lambda_i\} = \{\tilde{\lambda}_i + \sigma^2\} \quad i = 1, \ldots, p$$

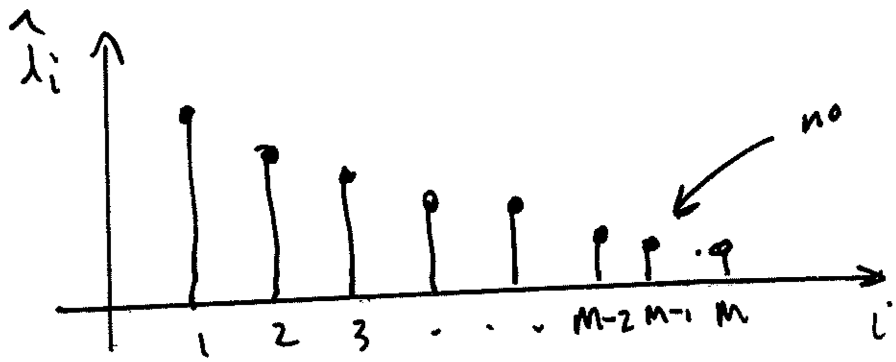
$$\{\lambda_i\} = \{\sigma^2\} \quad i = p+1, \ldots, M$$

Thus, if we order the eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_M$ we have



In this case, we have $p = M-3$.

Now, of course we only have the estimated covariance matrix $\underline{\hat{R}}$, and thus noisy estimates of the true eigenvalues.



How do we decide which correspond to signal and noise?

No sharp cut-off in general.

Minimum Description Length Criterion

Wax & Kailath (1985)

$$MDL(p) = -\log\left(\frac{r(p)}{q(p)}\right)^N + e(p)$$

$$r(p) = \prod_{i=p+1}^{M} \hat{\lambda}_i \qquad p = 0, 1, \ldots, M-1$$

$$q(p) = \left[\frac{1}{M-p}\sum_{i=p+1}^{M} \hat{\lambda}_i\right]^{M-p}$$

where $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \cdots \geq \hat{\lambda}_M$ are eigenvalues of $\hat{\underline{R}}$.

$$e(p) = \frac{1}{2}\, p\,(2M-p)\log N$$

$$\boxed{\hat{p} = \arg\min_p MDL(p)}$$

Interpretation:

$\frac{r(p)}{q(p)}$ measures how close the "tail" eigenvalues are to their average

$e(p)$ measures the "complexity" of using p sinusoidal components, given N measurements