CMPE 58P - Lecture 2.
Machine Listening

Sinusoidal Models, Estimation of Frequency

Department of Computer Engineering,

Boğaziçi University, Istanbul, Turkey

Instructor: A. Taylan Cemgil

Spring 2009

Outline

- Audio Representations, Sampling, PCM,
- Fourier transforms, FFT, Spectrogram, Short time Fourier transform,
- Sinusoidal models, Estimation of frequency

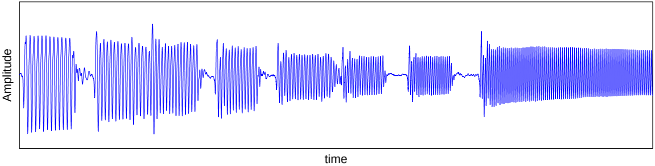
Useful Resources

- Freesound
  - Free samples of sounds: Creative Commons License
  - `http://www.freesound.org`
- LilyPond
  - Text based Music Notation Language
  - `lilypond.org`

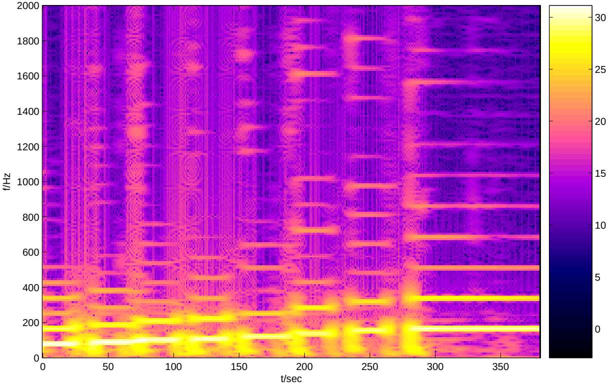
Digital Audio

A bunch of integers

$$\ldots\ 5386\ \ 5051\ \ 4693\ \ 4339\ \ 3966\ \ 3572\ \ 3162\ \ 2741\ \ldots$$



Digital Audio (Spectrogram)



Sinusoids

- Analog → ADC ⇢ Digital ⇢ DAC →
- Ideal sampling, PCM (Pulse Code Modulation)

$$x(nT) = \int x(t)\delta(t - nT_s)dt$$

- Sampling period T_s
- Sampling frequency $F_s = 1/T_s$
- Popular audio sampling frequencies 44100, 22050, 11025, 8000, 16000
  - $44100 = 2^2 3^2 5^2 7^2$
- Bits per sample: 24, 16, 8
- Channels: Mono (1), Stereo (2), Surround (5.1)

- Cosine and sine functions: Sinusoidal Signals or Sinusoids

$$x(t) = A\cos(\omega t + \phi)$$

Sinusoids (cont.)

- We will use typically $\cos$ because

$$
\begin{aligned}
\sin(x) &= \cos(x - \pi/2) \\
A\sin(\omega t + \hat{\phi}) &= A\cos(\omega t + \hat{\phi} - \pi/2)
\end{aligned}
$$

Sinusoids

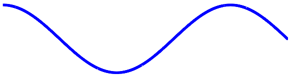
$$x(t) = A\cos(\omega t + \phi)$$

t : Time Index

A : Amplitude

ω : Angular velocity, radian frequency

ϕ : Phase



Sinusoids

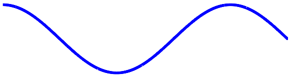
- Frequency

$$f = \frac{\omega}{2\pi}$$

- f is measured in Hertz, cycles per second
- ω is measured in radians per second

$$x(t) = A\cos(2\pi f t + \phi)$$

t : Time Index
A : Amplitude
f : (cyclic) frequency
ϕ : Phase



Periodicity

- A function is periodic if

$$x(t) = x(t + T_0)$$

- Sinusoids are periodic

$$\begin{aligned} A\cos(\omega t + \phi) &= A\cos(\omega(t + T_0) + \phi) \\ &= A\cos(\omega t + \omega T_0 + \phi) \end{aligned}$$

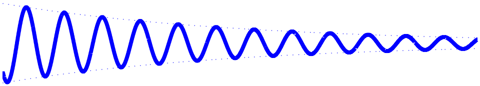
- The period is given by

$$\begin{aligned} \omega T_0 &= 2\pi \\ 2\pi f T_0 &= 2\pi \end{aligned}$$

$$\boxed{T_0 = \frac{1}{f}}$$

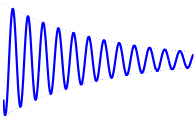
Damped Sinusoids

$$x(t) = Ae^{-\gamma t}\cos(\omega t + \phi) \quad (1)$$



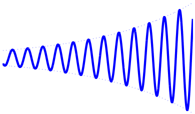
- 4 free parameters $\boxed{A, \gamma, \omega, \phi}$
- γ : negative log-damping coefficient
  - $0 < \exp(-\gamma) \leq 1$ when $\gamma > 0$

Damped Sinusoids



$\gamma > 0$

$\gamma = 0$

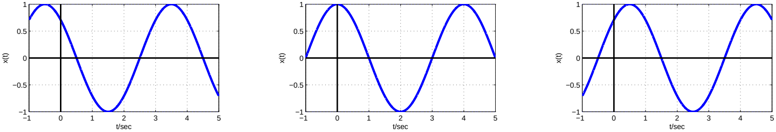


$\gamma < 0$

Phase Shift/Time Shift

- The phase ϕ determines locations of peaks

$$x(t) = A\cos(\omega t + \phi) \quad (2)$$

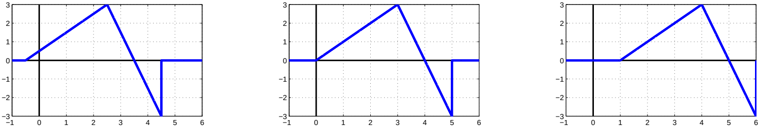


$\phi > 0$ $\phi = 0$ $\phi < 0$

Phase Shift/Time Shift

- This holds for all signals

$x(t + |\Delta|)$ $\quad$ $x(t)$ $\quad$ $x(t - |\Delta|)$



$\Delta > 0$ $\quad$ $\Delta = 0$ $\quad$ $\Delta < 0$

Positive phase shifts to **left**

Negative phase shifts to **right**

Complex Numbers (Cartesian Representation)

$$
\begin{aligned}
z &= x + jy \\
j &\equiv \sqrt{-1} \\
x &= \Re e\{z\} && \text{Real Part} \\
y &= \Im m\{z\} && \text{Imaginary Part} \\
z^* &= x - jy && \text{Conjugate}
\end{aligned}
$$

Polar Representation of Complex Numbers

$$z = x + jy$$

$$\boxed{z = re^{j\theta}}$$

$$r = \sqrt{x^2 + y^2}$$

$$\theta = \angle z = \arctan \frac{y}{x}$$

Euler's Formula

$$e^{j\theta} = \cos(\theta) + j\sin(\theta)$$

- "The most beautiful formula" in Mathematics

$$e^{-\pi j} + 1 = 0$$

Polar Representation

- Convenient

$$
\begin{aligned}
z^2 &= re^{j\theta} re^{j\theta} = r^2 e^{2j\theta} = e^{2(\log r j\theta)} \\
z^* &= re^{-j\theta} = e^{\log r - j\theta} \\
zz^* &= re^{j\theta} re^{-j\theta} = r^2
\end{aligned}
$$

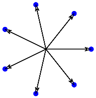
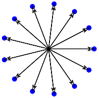
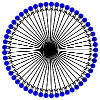
- Logarithms of complex numbers

$$
\begin{aligned}
z &= e^x \Leftrightarrow \log z = x \\
z &= e^{j\theta} \Leftrightarrow \log z = j\theta
\end{aligned}
$$

N'th root of unity

- All Solutions of the equation

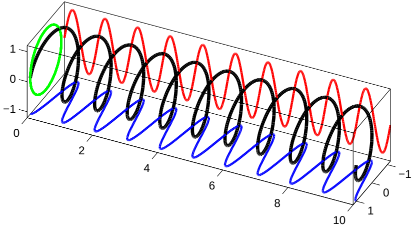
$$x^N = 1$$

$$\begin{array}{ll}
N = 1 & x = \{1\} \\
N = 2 & x = \{1, e^{j2\pi/2}\} = \{1, -1\} \\
N = 3 & x = \{1, e^{j2\pi/3}, e^{j4\pi/3}\} \\
N = 4 & x = \{1, e^{j2\pi/4}, e^{j4\pi/4}, e^{j6\pi/4}\} = \{1, j, -1, -j\} \\
\vdots & \vdots \\
N & x = \{e^{jk2\pi/N}\}_{k=0\ldots N-1}
\end{array}$$

Phasor Representation of Complex Exponential Signals

$$
\begin{aligned}
x(t) &= Ae^{j(\omega t+\phi)} = Ae^{j\phi}e^{j\omega t} = Xe^{j\omega t} \\
&= A\cos(\omega t+\phi) + jA\sin(\omega t+\phi)
\end{aligned}
$$



Rotation Matrices

- A (Given's) Rotation Matrix

$$R(\omega) = \begin{pmatrix} \cos(\omega) & -\sin(\omega) \\ \sin(\omega) & \cos(\omega) \end{pmatrix}$$

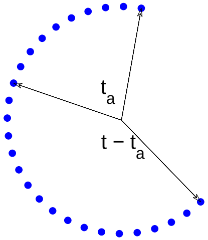
- An isomorphism

$$Xe^{j\omega} \iff R(\omega) \begin{pmatrix} \Re e\{X\} \\ \Im m\{X\} \end{pmatrix}$$

Complex Exponential Signals

$$x(t) = e^{j\omega(t-t_a)} e^{j\omega t_a} X$$

$$\begin{pmatrix} \Re e\{x(t)\} \\ \Im m\{x(t)\} \end{pmatrix} = R(\omega(t-t_a)) R(\omega t_a) \begin{pmatrix} \Re e\{X\} \\ \Im m\{X\} \end{pmatrix}$$

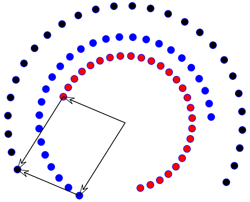


Phasor addition

- Sum of sinusoidals with the **same frequency** but **different phases** have the **same frequency**.

$$\sum_i A_i \cos(\omega t + \phi_i) = A \cos(\omega t + \phi)$$

- Easy to see on the complex plane via vector addition



Complex Exponential representations of Real sinusoids

- Phasor representation

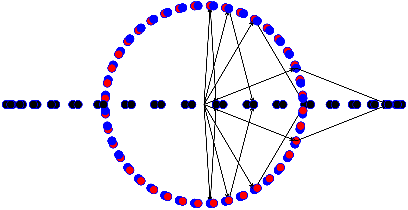
$$\sum_i A_i \cos(\omega t + \phi_i) = \sum_i \Re e\{A_i e^{j(\omega t + \phi_i)}\}$$
$$= \sum_i \Re e\{A_i e^{j\phi_i} e^{j\omega t}\}$$

- $\Re e\{\}$ is a linear operator

$$= \Re e\{e^{j\omega t} \sum_i A_i e^{j\phi_i}\} = \Re e\{e^{j\omega t} A e^{j\phi}\}$$
$$= A \cos(\omega t + \phi)$$

Complex Exponential representations of Real sinusoids

$$x(t) = Xe^{j\omega t} + X^* e^{-j\omega t}$$



Inverse Euler Formulas

$$\cos(\theta) = \frac{e^{j\theta} + e^{-j\theta}}{2} \qquad \sin(\theta) = \frac{e^{j\theta} - e^{-j\theta}}{2j}$$

Phasor representation of damped sinusoidals

By Euler's identity $\cos(\theta) = (e^{j\theta} + e^{-j\theta})/2$

$$
\begin{aligned}
x(t) &= Ae^{-\gamma t}\cos(\omega t + \phi) \\
&= \frac{A}{2} e^{-\gamma t}(e^{j\omega t + j\phi} + e^{-j\omega t - j\phi}) \\
&= \frac{A}{2}(e^{(-\gamma + j\omega)t} e^{j\phi} + e^{(-\gamma - j\omega)t} e^{-j\phi})
\end{aligned}
$$

Phasor representation of damped sinusoidals

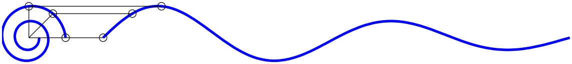
Define

$$c \equiv \frac{A}{2} e^{j\phi} \qquad \text{complex amplitudes}$$

$$z \equiv e^{-\gamma + j\omega} \qquad \text{complex poles}$$

Since $e^{(-\gamma + j\omega)t} = (e^{-\gamma + j\omega})^t$

$$x(t) = c z^t + c^* {z^*}^t$$



Complex poles and Amplitudes representation, vector notation

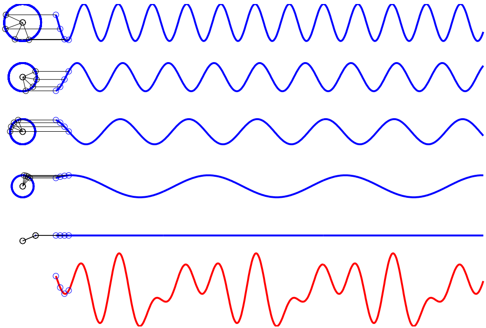
In compact vector notation, we denote $x_t = cz^t + c^* z^{*t}$

$$\mathbf{x} \equiv \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_t \\ \vdots \\ x_{N-1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ z & z^* \\ \vdots & \vdots \\ z^t & z^{*t} \\ \vdots & \vdots \\ z^{N-1} & z^{*N-1} \end{pmatrix} \begin{pmatrix} c \\ c^* \end{pmatrix}$$

- highlights the Vandermonde structure of the matrix

$\rightarrow$ good for efficient calculations

Spectrum Representations



$$x(t) = A_0 + \sum_{i=1}^{I} A_i \cos(\omega_i t + \phi_i) = X_0 + \Re e\{\sum_i X_i e^{j\omega_i t}\} \quad (3)$$

Spectrum Representations

$$x(t) = X_0 + \sum_i \left\{ \frac{X_i}{2} e^{j2\pi f_i t} + \frac{X_i^*}{2} e^{-j2\pi f_i t} \right\}$$

- Two sided spectrum

$$\{(0, X_0), (f_1, \frac{1}{2}X_1), (-f_1, \frac{1}{2}X_1^*), \dots (f_i, \frac{1}{2}X_i), (-f_i, \frac{1}{2}X_i^*), \dots\}$$

Complex Amplitude notation

$$c_i = \frac{1}{2} A_i e^{j\phi_i}$$

$$\begin{aligned} x(t) &= \sum_{i=0}^{I} c_i e^{j2\pi f_i t} + \sum_{i=0}^{I} c_i^* e^{-j2\pi f_i t} \\ &\equiv \sum_{i=0}^{I} \left(c_i z_i^t + c_i^* z_i^{*t} \right) \end{aligned}$$

Caution when $f_0 = 0$.

Complex Amplitude notation (alternative definition)

+

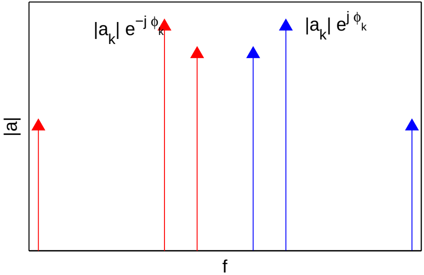
$$c_i = \begin{cases} A_i, & \text{for } f_i = 0 \\ A_i/2, & \text{for } f_i \neq 0 \end{cases}$$

$$x(t) = \sum_{i=-I}^{I} c_i e^{j2\pi f_i t} \equiv \sum_{i=-I}^{I} c_i z_i^t$$

We will always say which definition we use!

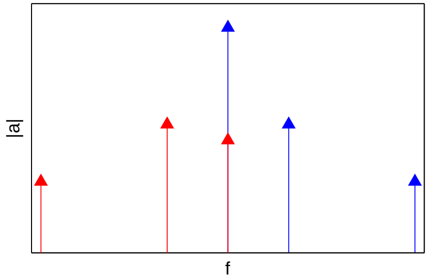
Graphic Representation of the Spectrum

$$x(t) \quad = \quad \sum_{i=0}^{I} c_i e^{j2\pi f_i t} + \sum_{i=0}^{I} c_i^* e^{-j2\pi f_i t}$$



Graphic Representation of the Spectrum ($f_0 = 0$)

$$x(t) = \sum_{i=0}^{I} c_i e^{j2\pi f_i t} + \sum_{i=0}^{I} c_i^* e^{-j2\pi f_i t}$$



Beating effect

- We know that sum of sinusoidals with the **same frequency** but **different phases** have the **same frequency.**

$$\sum_i A_i \cos(2\pi f t + \phi_i) = A \cos(2\pi f t + \phi)$$

- What happens if the frequencies f_i are close?

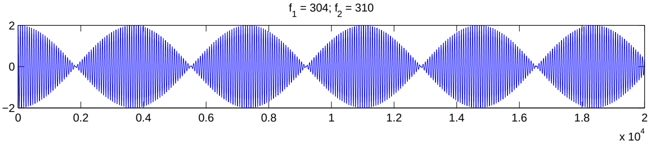
$$\sum_i A_i \cos(2\pi f_i t + \phi_i)$$

Example

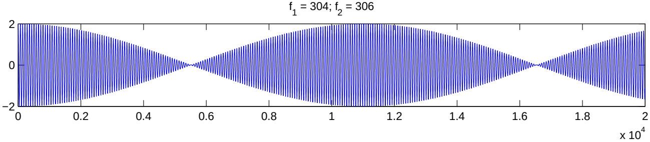
$$
\begin{aligned}
x(t) &= \cos(2\pi(f-\delta)t) + \cos(2\pi(f+\delta)t) \\
&= \frac{1}{2}\left(e^{j2\pi(f-\delta)t} + e^{-j2\pi(f-\delta)t}\right) + \frac{1}{2}\left(e^{j2\pi(f+\delta)t} + e^{-j2\pi(f+\delta)t}\right) \\
&= \frac{1}{2}\left(e^{j2\pi ft}e^{-j2\pi\delta t} + e^{-j2\pi ft}e^{j2\pi\delta t}\right) + \frac{1}{2}\left(e^{j2\pi ft}e^{j2\pi\delta t} + e^{-j2\pi ft}e^{-j2\pi\delta t}\right) \\
&= \frac{1}{2}\left(e^{j2\pi ft}e^{-j2\pi\delta t} + e^{-j2\pi ft}e^{j2\pi\delta t} + e^{j2\pi ft}e^{j2\pi\delta t} + e^{-j2\pi ft}e^{-j2\pi\delta t}\right) \\
&= \frac{1}{2}\left(\left(e^{j2\pi ft} + e^{-j2\pi ft}\right)e^{-j2\pi\delta t} + \left(e^{-j2\pi ft} + e^{j2\pi ft}\right)e^{j2\pi\delta t}\right) \\
&= \frac{1}{2}\left(e^{j2\pi ft} + e^{-j2\pi ft}\right)\left(e^{j2\pi\delta t} + e^{-j2\pi\delta t}\right)
\end{aligned}
$$

It's all about factorisation!

Beating effect

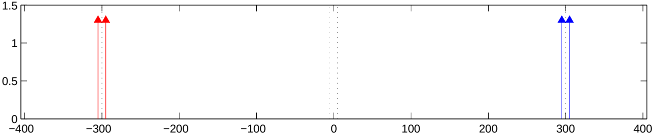


Beating effect



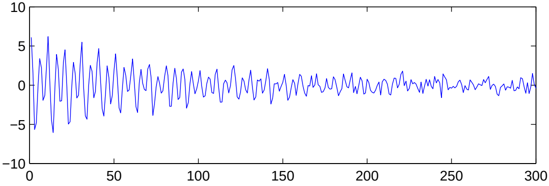
Beating effect

- Spectral plot of $A\cos(2\pi 300t)\cos(2\pi 5t)$



- "Multiplication in time domain is convolution in Frequency domain"
- When do we hear beating or two distinct frequencies ? Critical band

Sinusoidal Estimation Problem



- Given a sequence of samples $x_0, x_1, \ldots, x_{N-1}$, estimate
  - Frequencies
  - Phases
  - Amplitudes
  - Damping coefficients
  - Number of sinusoidals
- Backbone of any audio signal analysis task

Literature

- Rao and Arun, *Model based processing of signals, a state space approach*, Proc. of IEEE, 1992
- Sarkar and Pereira, *Using the matrix Pencil method to estimate the parameters of a sum of Complex exponentials*, IEEE Antennas and Propagation magazine, 1995
- Viberg, *Subspace-based Methods for the identification of Linear Time-invariant Systems*, Automatica, 1995
- Louis Sharf, *Statistical Signal Processing*, Ch. 11 (with Cedric Demeure), Addison Wesley, 1991

Motivations

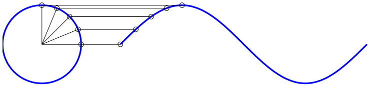
- Separation, Transcription, Restoration
- Getting inspiration from earlier techniques
  - ad-hoc and heuristic
  - Geometric intuition
  - Uses well understood tools from numerical linear algebra $\Rightarrow$ often faster and numerically stable

Sinusoidal model (without damping)

A pure sinusoidal for $t = 0 \dots N-1$ is given by

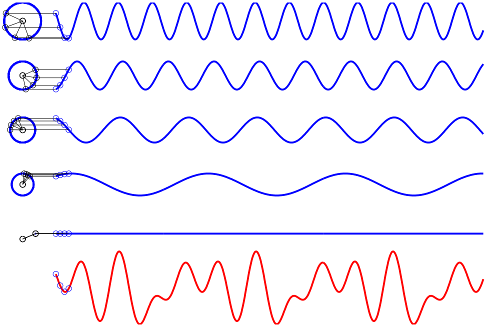
$$x_t = A\cos(\omega t + \phi) \tag{4}$$

F_s: sampling frequency
f : frequency $\qquad \omega \equiv 2\pi f / F_s$
A : amplitude $\qquad \phi$: phase



- 3 free parameters $\boxed{A, \omega, \phi}$.

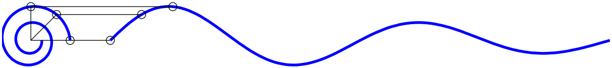
Sinusoidal model (without damping)



$$x_t = \sum_{i=1}^{I} A_i \cos(\omega_i t + \phi_i) \quad (5)$$

Sinusoidal model (with exponential damping)

$$x_t \quad = \quad Ae^{-\gamma t}\cos(\omega t + \phi) \tag{6}$$



- 4 free parameters $\boxed{A, \gamma, \omega, \phi}$
- γ : negative log-damping coefficient
  - ⋆ $0 < \exp(-\gamma) \leq 1$ when $\gamma > 0$, good for positive priors

Complex poles and Amplitudes representation of a damped sinusoidal

By using basic trigonometric identity $\cos(\theta) = (e^{j\theta} + e^{-j\theta})/2$

$$
\begin{aligned}
x_t &= Ae^{-\gamma t}\cos(\omega t + \phi) \\
&= \frac{A}{2}e^{-\gamma t}(e^{j\omega t + j\phi} + e^{-j\omega t - j\phi}) \\
&= \frac{A}{2}(e^{(-\gamma + j\omega)t}e^{j\phi} + e^{(-\gamma - j\omega)n}e^{-j\phi})
\end{aligned}
$$

Define

$$
\begin{aligned}
c &\equiv \frac{A}{2}e^{j\phi} && \text{complex amplitudes} \\
z &\equiv e^{-\gamma + j\omega} && \text{complex poles}
\end{aligned}
$$

Let $*$ denote complex conjugate. Since $e^{(-\gamma + j\omega)t} = (e^{-\gamma + j\omega})^t$

$$
x_t = cz^t + c^* z^{*t}
$$

Complex poles and Amplitudes representation, vector notation

In compact vector notation, we denote $x_t = cz^t + c^* z^{*t}$

$$\mathbf{x} \equiv \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_t \\ \vdots \\ x_{N-1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ z & z^* \\ \vdots & \vdots \\ z^t & z^{*t} \\ \vdots & \vdots \\ z^{N-1} & z^{*N-1} \end{pmatrix} \begin{pmatrix} c \\ c^* \end{pmatrix}$$

- highlights the Vandermonde structure of the matrix

$\rightarrow$ good for efficient calculations

Sinusoidal model (with damping)

$$x_t = \sum_{i=1}^{I} c_i z_i^t + c_i^* z_i^{*t}$$

$$\mathbf{x} = \begin{pmatrix} 1 & \dots & 1 & 1 & \dots & 1 \\ z_1 & \dots & z_I & z_1^* & \dots & z_I^* \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ z_1^n & \dots & z_I^t & z_1^{*t} & \dots & z_I^{*t} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ z_1^{N-1} & \dots & z_I^{N-1} & z_1^{*N-1} & \dots & z_I^{*N-1} \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_I \\ c_1^* \\ \vdots \\ c_I^* \end{pmatrix}$$

Note the equivalence to the inverse Fourier transform when $\gamma = 0$ and $\omega_0 = 2\pi/N$.

Prototypical Problem: Sinusoidal detection

- Determine (complex) amplitudes c_i and frequencies ω_i for $i = 1 \ldots I$

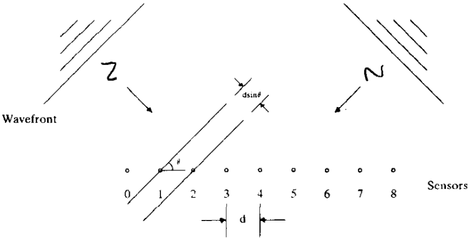
$$x_t = \sum_{i=1}^{I} c_i e^{j\omega_i t}$$

Direction of Arrival (DOA) estimation

- θ : Direction of arrival, d : Spacing between sensors (assumed to be uniform), v_c : Velocity of propagation
- Object index $i = 1 \dots I$, number of incident planewaves,
- Sensor index $i = 1 \dots L$, number of sensors on a linear array

$$x_{t,i} = \sum_{i=1}^{I} s_i(t - i\tau_i) \qquad \tau_i = \frac{d \sin \theta_i}{v_c}$$

Direction of Arrival (DOA) estimation



In frequency domain

$$x_{t,i} = \sum_{i=1}^{I} s_i(t - i\tau_i) \qquad \tau_i = \frac{d \sin \theta_i}{v_c}$$

$$Y_i(\nu) = \sum_{i=1}^{I} S_i(\nu) e^{-j\nu\tau_i i}$$

- At a fixed frequency band, $\hat{\nu}$, the output of the array element i is

$$\hat{Y}_i = \sum_{i=1}^{I} S_i e^{-j\omega_i i} \qquad \omega_i = \hat{\nu}\tau_i$$

- Compare this with the sinusoidal detection problem where i is the new “time” index and S_i is the complex amplitude

Direct Approach

- Minimize directly

$$(c_{1:I}, \omega_{1:I})^* = \arg \min_{c_{1:I}, \omega_{1:I}} \left| x_t - \sum_i c_i e^{j\omega_i t} \right|^2$$

- A potentially hard objective function with a lot of local minima
- Some analytic structure (conditioned on ω, c can be found easily), but
  - Computationally expensive
  - Essentially viable only in very low SNR
- More or less related to what we are trying to do in direct Bayesian inference

Indirect Approaches, Topic of this talk

- Fit a model with parameters linearly related to the signal
- Extract desired quantities from estimated parameters
- A rough classification
  - Polynomial
  - SVD (Singular Value Decomposition) or EVD (Eigenvalue Decomposition)
  - Subspace methods

Linear Models

- Sinusoidals can be generated by zero input oscillators
- The transfer function has poles on the unit circle
- In the absence of noise, an exponential signal satisfies exactly the relation

$$x_t = \sum_{i=1}^{I} a_i x_{t-i}$$

State Space view, Companion or Canonical form

$$s^{\mathsf{c}}_t = \begin{pmatrix} x_{t-1} \\ x_{t-2} \\ \vdots \\ x_{t-I} \end{pmatrix}$$

$$s^{\mathsf{c}}_{t+1} = \underbrace{\begin{pmatrix} a_1 & a_2 & \dots & a_{I-1} & a_I \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix}}_{A_{\mathsf{c}}} s^{\mathsf{c}}_t$$

$$x_t = \underbrace{\begin{pmatrix} a_1 & a_2 & \dots & a_{I-1} & a_I \end{pmatrix}}_{C_{\mathsf{c}}} s^{\mathsf{c}}_t$$

Diagonal Parametrisation

$$s_0^{\mathsf{d}} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_I \end{pmatrix}$$

$$s_{t+1}^{\mathsf{d}} = \underbrace{\begin{pmatrix} e^{j\omega_1} & & \\ & \ddots & \\ & & e^{j\omega_I} \end{pmatrix}}_{A_{\mathsf{d}}} s_t^{\mathsf{d}}$$

$$x_t = \underbrace{\begin{pmatrix} 1 & 1 & \dots & 1 & 1 \end{pmatrix}}_{C_{\mathsf{d}}} s_t^{\mathsf{d}}$$

Diagonal Parametrisation for strictly real x

$$s_0^{\mathsf{dr}} = \begin{pmatrix} c_1 \\ c_1^* \\ \vdots \\ c_I^* \\ c_I \end{pmatrix}$$

$$s_{t+1}^{\mathsf{dr}} = \underbrace{\begin{pmatrix} e^{j\omega_1} & & & & \\ & e^{-j\omega_1} & & & \\ & & \ddots & & \\ & & & e^{j\omega_I} & \\ & & & & e^{-j\omega_I} \end{pmatrix}}_{A_{\mathsf{dr}}} s_t^{\mathsf{dr}}$$

$$x_t = \underbrace{\begin{pmatrix} 1 & 1 & \dots & 1 & 1 \end{pmatrix}}_{C_{\mathsf{dr}}} s_t^{\mathsf{dr}}$$

Alternative realizations

- A linear system has infinitely many relalisations each corresponding to a state space representation

$$\bar{s}_{t+1} = \bar{A}\bar{s}_t$$
$$x_t = \bar{C}\bar{s}_t$$

with

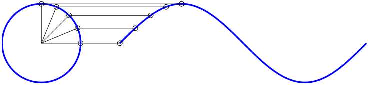
$$\bar{s}_t \equiv Ts_t \qquad \bar{A} \equiv TAT^{-1} \qquad \bar{C} \equiv CT^{-1}$$

$$Ts_{t+1} = TAT^{-1}Ts_t$$
$$x_t = CT^{-1}Ts_t$$

- There are many parametrisations, but some are more suitable for extracting relevant information about a system, such as poles.

Example: Block Diagonal Parametrisation for strictly real x

$\cos(\theta) = \frac{1}{2}(e^{j\theta} + e^{-j\theta})$



$\sin(\theta) = j\frac{1}{2}$

$$T \equiv \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -j & j \end{pmatrix} \qquad TT^H = T^H T = I$$

$$A^{\mathsf{b}} = TA^{\mathsf{d}}T^H = T \begin{pmatrix} e^{j\omega} & 0 \\ 0 & e^{-j\omega} \end{pmatrix} T^H = \begin{pmatrix} \cos(\omega) & -\sin(\omega) \\ \sin(\omega) & \cos(\omega) \end{pmatrix}$$

$$s_0^{\mathsf{b}} = Ts_0^{\mathsf{d}} = T \begin{pmatrix} c \\ c^* \end{pmatrix} = \sqrt{2} \begin{pmatrix} \mathsf{Re}\{c\} \\ \mathsf{Im}\{c\} \end{pmatrix}$$

$$C^{\mathsf{b}} = C^{\mathsf{d}}T^H = \begin{pmatrix} 1 & 1 \end{pmatrix} T^H = \sqrt{2} \begin{pmatrix} 1 & 0 \end{pmatrix}$$

Rank Properties

- The state space representation implies

$$\begin{array}{rcl} s_{t+1} & = & A s_t \\ x_t & = & C s_t \end{array} \quad \Rightarrow \quad x_t = C A^t s_0$$

- Therefore we can write for arbitrary L and M the Hankel matrix

$$\underbrace{\begin{pmatrix} y_0 & y_1 & \dots & y_M \\ y_1 & y_2 & \dots & y_{M+1} \\ \vdots & \vdots & \ddots & \vdots \\ y_L & y_{L+1} & \dots & y_{L+M} \end{pmatrix}}_{Y} = \underbrace{\begin{pmatrix} C \\ CA \\ \vdots \\ CA^L \end{pmatrix}}_{\Gamma_{L+1}} \underbrace{\begin{pmatrix} s_0 & As_0 & \dots & A^M s_0 \end{pmatrix}}_{\Omega_{M+1}}$$

- When $L \geq I$ and $M \geq I$, $\mathbf{rank}(\Gamma) = \mathbf{rank}(\Omega) = I = \mathbf{rank}(Y)$.

Identification via matrix factorisation (Ho and Kalman 1966)

- Given the data matrix Y layout as a Hankel matrix
  1. Compute a factorisation (e.g., via SVD)

$$Y = \bar{\Gamma}_{L+1}\bar{\Omega}_{M+1} = \begin{pmatrix} C \\ CA \\ \vdots \\ CA^L \end{pmatrix} \begin{pmatrix} s_0 & As_0 & \dots & A^M s_0 \end{pmatrix}$$

  2. The resulting factors $\bar{\Gamma}_{L+1}\bar{\Omega}_{M+1}$ correspond to observability and controllability matrices of *some* realisation
  3. Read off C and s_0 from factors $\bar{\Gamma}_{L+1}$ and $\bar{\Omega}_{M+1}$
  4. Compute transition matrix by exploiting *shift invariance*

$$\begin{pmatrix} CA \\ CA^2 \\ \vdots \\ CA^L \end{pmatrix} = \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{L-1} \end{pmatrix} A \quad \Rightarrow \quad A = \Gamma_{1:L}^{\dagger}\Gamma_{2:L+1}$$

Remarks

- When A is diagonal, Γ is a Vandermonde matrix with $z_i = e^{j\omega_i}$

$$\Gamma_{L+1} = \begin{pmatrix} 1 & \dots & 1 \\ z_1 & \dots & z_I \\ \vdots & \vdots & \vdots \\ z_1^n & \dots & z_I^t \\ \vdots & \vdots & \vdots \\ z_1^L & \dots & z_I^L \end{pmatrix}$$

- Matrices of form

$$\Omega_{M+1} = \begin{pmatrix} s_0 & As_0 & \dots & A^M s_0 \end{pmatrix}$$

are known as *Krylov* matrices (Golub and Van Loan p.p. 347)

Singular Value Decomposition (SVD)

For any $A \in \mathbb{R}^{m \times n}$, there exist **orthogonal** matrices $U \in \mathbb{R}^{m \times m}$ and $V \in \mathbb{R}^{n \times n}$ such that

$$U = [u_1, \ldots, u_m] \qquad\qquad V = [v_1, \ldots, v_n]$$

such that

$$U^\top A V = \mathbf{diag}(\sigma_1, \ldots, \sigma_p) \in \mathbb{R}^{m \times n}$$

with $p = \min\{m, n\}$. We have

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p \geq 0$$

Singular Value Decomposition (SVD)

```
>> A = randn(3,2); [U S V] = svd(A)
U =
   -0.1336    0.8386   -0.5280
   -0.4548   -0.5253   -0.7192
   -0.8805    0.1440    0.4516
S =
    2.5216         0
         0    1.1498
         0         0
V =
    0.9973    0.0731
   -0.0731    0.9973
```

Singular Value Decomposition (SVD)

The σ_i are the *singular values*, u_i and v_i are the left singular vectors and right singular vectors, respectively. We have the following relationships:

$$\begin{aligned} Av_i &= \sigma_i u_i \\ u_i^\top A &= \sigma_i v_i^\top \end{aligned}$$

Geometrically, the singular values are the lengths of the semi-axes of the hyperellipsoid E obtained by transforming the unit sphere via A, i.e.,

$$E \equiv \{Ax : \|x\|_2 = 1\}$$

Singular Value Decomposition (SVD)

SVD tells about the structure of a matrix. If

$$\sigma_1 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_p = 0$$

$$\begin{aligned} \textbf{rank}(A) &= r \\ \textbf{null}(A) &= \textbf{span}(v_{r+1}, \ldots, v_n) \\ \textbf{range}(A) &= \textbf{span}(u_1, \ldots, u_r) \end{aligned}$$

Singular Value Decomposition (SVD)

SVD expansion

$$A = \sum_{i}^{r} \sigma_i u_i v_i^\top$$

The norm relations for $A \in \mathbb{R}^{m \times n}$, $p = \min\{m, n\}$

$$\begin{aligned} \|A\|_F^2 &= \sigma_1^2 + \cdots + \sigma_p^2 \\ \|A\|_2^2 &= \sigma_1^2 \end{aligned}$$

See Cleve Moler, “Professor SVD”, The Mathworks News and Notes, October 2006

(Moore-Penrose) Pseudo-inverse

Let $A \in \mathbb{R}^{m \times n}$ with $\mathbf{rank}(A) = r$. The pseudo-Inverse $A^\dagger$ is defined via SVD as follows:

$$\begin{aligned} A^\dagger &= V\Sigma^\dagger U^\top \\ \Sigma^\dagger &\equiv \mathbf{diag}(\frac{1}{\sigma_1}, \dots, \frac{1}{\sigma_r}, 0, \dots, 0) \end{aligned}$$

The pseudo inverse is important for rank defficient least-square problems where $r < \min\{m, n\}$.

$$\begin{aligned} Ax &= b \\ x_{LS} &= A^\dagger b \end{aligned}$$

(Moore-Penrose) Pseudo-inverse

$A^\dagger$ is the unique solution for

$$A^\dagger = \underset{X \in \mathbb{R}^{n\times m}}{\operatorname{argmin}} \|AX - I_m\|_F$$

- If $\mathbf{rank}(A) = n$, then $A^\dagger = (A^\top A)^{-1} A^\top$.
- If $m = n = \mathbf{rank}(A)$, $A^\dagger = A^{-1}$.

(Moore-Penrose) Pseudo-inverse

- $A^\dagger$ is the unique matrix $X \in \mathbb{R}^{n \times m}$ that satisfies

$$AXA = A \qquad (AX)^\top = AX$$
$$XAX = X \qquad (XA)^\top = XA$$

This is equivalent to $AA^\dagger$ and $A^\dagger A$ being orthogonal projections onto $\mathbf{range}(A)$ and $\mathbf{range}(A^T)$. Indeed,

$$\begin{aligned} AA^\dagger &= U(:, 1:r)U(:, 1:r)^\top \\ A^\dagger A &= V(:, 1:r)V(:, 1:r)^\top \end{aligned}$$

Matlab Code

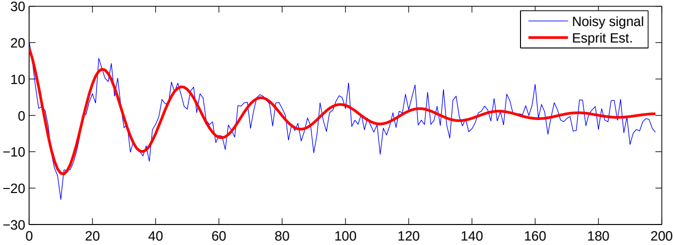
```
function [f, a, phi, rho] =  sub_esprit_svd(y, M)

% Length of the signal has
%  to be odd to setup the Hankel matrix
N = length(y); L = floor((N-1)/2); N = 2*L - 1;

% Setup the Hankel matrix and compute a facorisatio
H = hankel(y(1:L),y(L:N));
[U_full S V] = svd(H);
U = U_full(:, 1:2*M);

% Compute the transition matrix
U_t = U(1:end-1,:);
U_b = U(2:end,:);
A = pinv(U_t)*U_b;
```

Matlab Code

```
% Compute the diagonal realisation
% and read off the eigenvalues
[C, D] = eig(A); lambda = diag(D);

% Frequencies and log Damping factors
f = angle(lambda)/(2*pi);
rho = log(abs(lambda));

% Construct the Vandermonde matrix of poles
pows = repmat((0:L-1)',[1 length(f)]);
Z = repmat(lambda.', [L 1]).^pows;

% Compute the complex amplitudes
c = Z\y(1:L);
% Amplitudes and phases
a = abs(c); phi = angle(c);
```

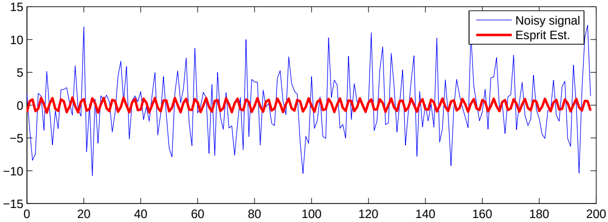
ESPRIT: Some results. $M = 1$, $M^{\text{true}} = 1$.

A damped sinusoid present



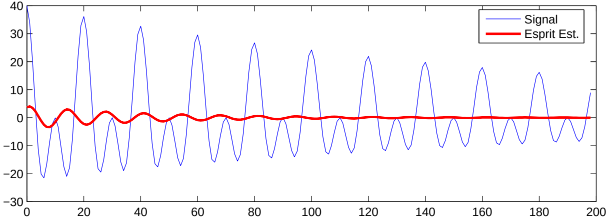
ESPRIT: $M = 1$, $M^{\text{true}} = 0$.

White noise

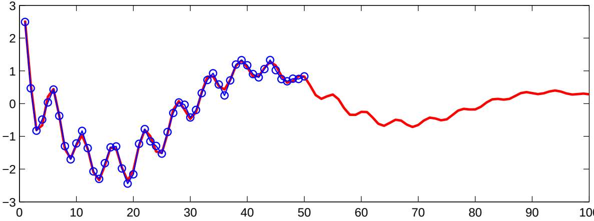


ESPRIT: $M = 1$, $M^{\text{true}} = 2$.

Two sinusoidals

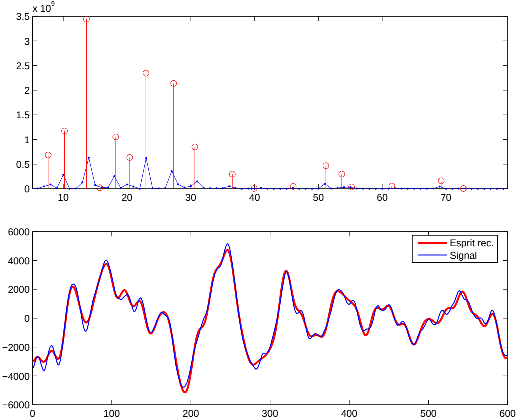


ESPRIT: $M = 8$, $M^{\text{true}} = 3$.



- Correct model order selection is a key issue !
  - Under estimation leads to **loss of accuracy**
  - Over estimation leads to **loss of computation time**

Real Data


x 10^9
3.5
3
2.5
2
1.5
1
0.5
0
10
20
30
40
50
60
70
6000
4000
2000
0
−2000
−4000
−6000
0
100
200
300
400
500
600
Esprit rec.
Signal

Appendix, Alternative Methods

Forward and Backward prediction for sinusoids

- For the diagonal realisation we have $A_{\mathsf{d}}^{-1} = A_{\mathsf{d}}^{H} = A_{\mathsf{d}}^{*}$

$$
\begin{aligned}
x_t &= C_{\mathsf{d}} A_{\mathsf{d}}^{t} s_0^{\mathsf{d}} && \text{forward} \\
&= C_{\mathsf{d}} \left(A_{\mathsf{d}}^{-1}\right)^{N-t} A_{\mathsf{d}}^{N} s_0^{\mathsf{d}} = C_{\mathsf{d}} \left(A_{\mathsf{d}}^{*}\right)^{N-t} s_N^{\mathsf{d}} \\
x_t^{*} &= C_{\mathsf{d}} A_{\mathsf{d}}^{N-t} s_N^{\mathsf{d}*} && \text{backward}
\end{aligned}
$$

- For an invertible transformation matrix T

$$
\begin{aligned}
x_t^{*} &= C_{\mathsf{d}} T^{-1} T A_{\mathsf{d}}^{N-t} T^{-1} T s_N^{\mathsf{d}*} \\
&= \bar{C} \bar{A}^{N-t} \bar{s}_N
\end{aligned}
$$

- By Cayley-Hamilton theorem $A_{\mathsf{c}}^{t} = \sum_{i=1}^{I} a_i A_{\mathsf{c}}^{t-i}$

$$
x_t^{*} = \bar{C} \bar{A}^{N-t} \bar{s}_N = \sum_{i=1}^{I} a_i \bar{C} \bar{A}^{N-t-i} \bar{s}_N = \sum_{i=1}^{I} a_i y^{*}(t+i)
$$

Covariances

- Consider the m-lag correlations r_m with respect to a fixed time index t

$$
\begin{aligned}
r_m &= \langle x_{t+m} x_t^* \rangle \\
x_{t+m} &= C A^m s_t \qquad\qquad x_t^* = s_t^H C^H \\
r_m &= \left\langle C A^m s_t s_t^H C^H \right\rangle = C A^m \left\langle s_t s_t^H \right\rangle C^H \equiv C A^m P_t C^H \\
r_{-m} &= C A^{-m} P_t C^H
\end{aligned}
$$

Covariances

- The toeplitz cross correlation matrix R can be factorised as

$$\underbrace{\begin{pmatrix} r_0 & r_{-1} & r_{-2} & \dots & r_{-t} \\ r_1 & r_0 & r_{-1} & \dots & r_{-t+1} \\ r_2 & r_1 & r_0 & \dots & r_{-t+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_t & r_{t-1} & r_{t-2} & \dots & r_0 \end{pmatrix}}_{R_{t+1}} = \underbrace{\begin{pmatrix} C \\ CA \\ \vdots \\ CA^t \end{pmatrix}}_{\Gamma_{t+1}} \Big\langle \underbrace{\begin{pmatrix} s_0 & A^{-1}s_0 & \dots & A^{-t}s_0 \end{pmatrix}}_{\Theta_{t+1}}$$

$$R_{t+1} = \begin{pmatrix} C \\ CA \\ \vdots \\ CA^t \end{pmatrix} \begin{pmatrix} P_0C^H & A^{-1}P_0C^H & \dots & A^{-t}P_0C^H \end{pmatrix}$$

Prony's method (1795)

- Exploits the fact that $x_t = \sum_{i=1}^{I} a_i x_{t-i}$ (or $x_t^* = \sum_{i=1}^{I} a_i x_{t+i}^*$)
- Set up a set of linear equations and compute a_i for $i = 1 \dots I$

$$\begin{pmatrix} x_{I-1} & x_{I-2} & \dots & x_0 \\ x_I & x_{I-1} & \dots & x_1 \\ \dots & & & \\ x_{N-2} & x_{N-3} & \dots & x_{N-I-1} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_I \end{pmatrix} = \begin{pmatrix} x_I \\ x_{I+1} \\ \vdots \\ x_{N-1} \end{pmatrix}$$

- Desired frequencies are the arguments of the complex roots of

$$H(z) = 1 - \sum_{i=1}^{I} a_i z^{-i}$$

which can be estimated by ploting $1/\left|H(e^{j\omega})\right|^2$ (spectral-approach) or finding the canonical factorisation of the polynomial (root-approach)

Prony's method - variations

- Forward and Backward method

$$\underbrace{\begin{pmatrix} x_{I-1} & x_{I-2} & \dots & x_0 \\ x_I & x_{I-1} & \dots & x_1 \\ \dots & & & \\ x_{N-2} & x_{N-3} & \dots & x_{N-I-1} \\ \hline x^*_{N-I} & x^*_{N-I-1} & \dots & x^*_{N-1} \\ x^*_{N-I-1} & x^*_{N-I-2} & \dots & x^*_{N-2} \\ \dots & & & \\ x^*_1 & x^*_2 & \dots & x^*_I \end{pmatrix}}_{Y} \underbrace{\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_I \end{pmatrix}}_{a} = \underbrace{\begin{pmatrix} x_I \\ x_{I+1} \\ \vdots \\ x_{N-1} \\ \hline x^*_{N-I-1} \\ x^*_{N-I-2} \\ \vdots \\ x^*_0 \end{pmatrix}}_{b}$$

$$\begin{pmatrix} b & Y \end{pmatrix} \begin{pmatrix} -1 \\ a \end{pmatrix} = 0$$

- $\begin{pmatrix} -1 \\ a \end{pmatrix}$ is in the null space of $\begin{pmatrix} b & Y \end{pmatrix}$

Prony's method - variations in finding a

- by Least Squares

$$a_{\text{LS}} = Y^{\dagger} b = (Y^H Y)^{-1} Y^H b$$

This implicitely assumes that noise is coming entirely from b

- by Total Least Squares

$$\min \| \begin{pmatrix} \delta & \Delta \end{pmatrix} \|_F \quad s.t \quad (Y - \Delta) a = (b - \delta)$$

  - i.e., finding a with the smallest perturbation to Y and b such that we obtain a solution
  - Can be obtained directly from the SVD $\begin{pmatrix} b & Y \end{pmatrix} = U \Sigma V^H$ where v is the right singular vector corresponding to σ^{min}

$$\begin{pmatrix} -1 \\ a_{TLS} \end{pmatrix} = -v/v_1$$

Pisarenko's method

- When the signal is pure sinusoidal, the cross correlations satisfy the linear prediction equation

$$\begin{pmatrix} r_0 & r_{-1} & r_{-2} & \dots & r_{-I} \\ r_1 & r_0 & r_{-1} & \dots & r_{-I+1} \\ r_2 & r_1 & r_0 & \dots & r_{-I+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_I & r_{I-1} & r_{I-2} & \dots & r_0 \end{pmatrix} \begin{pmatrix} -1 \\ a_1 \\ a_2 \\ \vdots \\ a_I \end{pmatrix} = 0$$

$$R_{I+1} \begin{pmatrix} -1 \\ a \end{pmatrix} = 0$$

- When the additive noise is white, only r_0 is affected

Pisarenko's method (cont.)

- Usually the method is described as extracting the prediction polynomial coefficients from the eigenvector corresponding to the smallest eigenvalue of R_{I+1}

$$R_{I+1} = \begin{pmatrix} U & u_{I+1} \end{pmatrix} \begin{pmatrix} \Lambda & \\ & \lambda_{I+1} \end{pmatrix} \begin{pmatrix} U^H \\ u_{I+1}^H \end{pmatrix}$$

- In noise-free case, $\lambda_{I+1} = 0$ because $\mathbf{rank}(R_{I+1}) = I$, i.e.,

$$R_{I+1} u_{I+1} = \lambda_{I+1} u_{I+1} = 0$$

- When $Y^H Y$ is used to calculate the cross correlation matrix R, Pisarenko is the same as total least squares Prony
- Restatement of Caratheodory's theorem
  Given complex numbers r_t for $t = 0 \dots L-1$, there exist unique real A_i and ω_i, $i = 1 \dots I$ when I is minimised

$$r_t = \sum_{i=1}^{I} A_i e^{j\omega_i t}$$

Conceptual and Practical Problems

- The polynomial coefficients are estimated by a vector that lies in the null space, i.e., so called *noise subspace*.
- Since point estimates are used, the errors in the polynomial coefficients transfer in a rather complicated way to the frequency estimate obtained from the roots.
- In practice, the correct number of frequencies I is not known and one usually chooses a frame length $L \geq I$. The "noise subspace" is typically identified as corresponding to the smallest $L + 1 - p$ eigenvalues.
- The handling of uncertainty is rather ad-hoc!