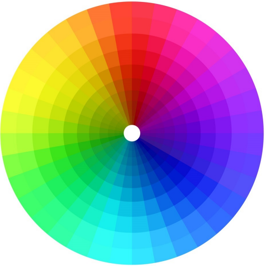
UPPSALA
UNIVERSITET

Subspace methods for line spectra



Per Mattsson

Systems and Control
Department of Information Technology
Uppsala University

2019-09-19

Summary from last lecture

- Signal model for line spectra
- Estimating the spectrum $\Longrightarrow$ frequency estimation
- Nonlinear least squares
- Models of the covariance matrix and ARMA-type model
- Singular value decomposition (SVD)

Today

- A recap from L5
- Subspace methods for frequency estimation
  - High order Yule-Walker
  - MUSIC
  - Min-Norm
  - ESPRIT
- Makes use of the covariance structure

Line spectra

Signal model:

$$\boxed{y(t) = \sum_{k=1}^{n} \alpha_k e^{i(\omega_k t + \varphi_k)} + e(t)}$$

Line spectra

Signal model:

$$\boxed{y(t) = \sum_{k=1}^{n} \alpha_k e^{i(\omega_k t + \varphi_k)} + e(t)}$$

Theoretical ACS for line spectra

$$\boxed{r(k) = E\{y(t)y^*(t-k)\} = \sum_{p=1}^{n} \alpha_p^2 e^{i\omega_p k} + \sigma^2 \delta_{k,0}}$$

per.mattssonk@it.uu.se

Line spectra

Signal model:

$$\boxed{y(t) = \sum_{k=1}^{n} \alpha_k e^{i(\omega_k t + \varphi_k)} + e(t)}$$

Theoretical ACS for line spectra

$$\boxed{r(k) = E\{y(t)y^*(t-k)\} = \sum_{p=1}^{n} \alpha_p^2 e^{i\omega_p k} + \sigma^2 \delta_{k,0}}$$

Theoretical spectrum

$$\boxed{\phi(\omega) = 2\pi \sum_{p=1}^{n} \alpha_p^2 \delta(\omega - \omega_p) + \sigma^2}$$

Frequency estimation

Problem

Determine $\{\omega_k\}_{k=1}^{n}$ from observations or samples $\{y(t)\}_{t=1}^{N}$

Frequency estimation

Problem

Determine $\{\omega_k\}_{k=1}^{n}$ from observations or samples $\{y(t)\}_{t=1}^{N}$

- The nonlinear least square method,

$$\min_{\omega,\alpha,\varphi} \sum_{k=1}^{N} \left| y(t) - \sum_{k=1}^{n} \alpha_k e^{i(\omega_k t + \varphi_k)} \right|^2$$

Frequency estimation

Problem

Determine $\{\omega_k\}_{k=1}^{n}$ from observations or samples $\{y(t)\}_{t=1}^{N}$

- The nonlinear least square method,

$$\min_{\omega,\alpha,\varphi} \sum_{k=1}^{N} \left| y(t) - \sum_{k=1}^{n} \alpha_k e^{i(\omega_k t + \varphi_k)} \right|^2$$

- Excellent properties if the global minimum can be found.

per.mattssonk@it.uu.se

Frequency estimation

Problem

Determine $\{\omega_k\}_{k=1}^{n}$ from observations or samples $\{y(t)\}_{t=1}^{N}$

- The nonlinear least square method,

$$\min_{\omega,\alpha,\varphi} \sum_{k=1}^{N} \left| y(t) - \sum_{k=1}^{n} \alpha_k e^{i(\omega_k t + \varphi_k)} \right|^2$$

- Excellent properties if the global minimum can be found.
- Hard (impossible) to use in general.

Frequency estimation

Problem

Determine $\{\omega_k\}_{k=1}^{n}$ from observations or samples $\{y(t)\}_{t=1}^{N}$

- The nonlinear least square method,

$$\min_{\omega,\alpha,\varphi} \sum_{k=1}^{N} \left| y(t) - \sum_{k=1}^{n} \alpha_k e^{i(\omega_k t + \varphi_k)} \right|^2$$

- Excellent properties if the global minimum can be found.
- Hard (impossible) to use in general.
- Today we will instead use subspace methods.

UPPSALA UNIVERSITET

Singular value decomposition (SVD)

Definition

Factor a matrix $A \in \mathbb{C}^{m \times n}$ as

$$A = U \Sigma V^*$$

where $U \in \mathbb{C}^{m \times m}$ and $V \in \mathbb{C}^{n \times n}$ are unitary matrices, and $\Sigma \in \mathbb{R}^{m \times n}$ is diagonal with non-negative diagonal values σ_i (singular values) ordered in decreasing order

(See A.4 in the book for the details)

- Low-rank approximations
- Useful to find the best fit subspaces of A
- Projection onto subspace – perpendicular distance
- LS fit – vertical distance

SVD (2)

$A \in \mathbb{C}^{m\times n}$ is of rank $r < \min(m,n)$ then

$$A = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^* \\ V_2^* \end{bmatrix} = U_1 \Sigma_1 V_1^*$$

where $\Sigma_1 \in \mathbb{R}^{r\times r}$, then

- U_1 is an orthonormal basis in $\mathcal{R}(A)$
- U_2 is an orthonormal basis in $\mathcal{N}(A^*)$
- V_1 is an orthonormal basis in $\mathcal{R}(A^*)$
- V_2 is an orthonormal basis in $\mathcal{N}(A)$

Moreover, $\mathcal{R}(A)$ and $\mathcal{N}(A^*)$ are orthogonal to each other and they span $\mathbb{C}^m$
($\mathcal{N}(A^*)$ is the orthogonal complement to $\mathcal{R}(A)$ and vice versa)

High-order Yule-Walker (HOYW)

High-order Yule-Walker (HOYW)

- Start from ARMA-model:

$$A(z)y(t) = A(z)e(t), \quad A(z) = \prod_{k=1}^{n}(1 - e^{i\omega_k} z^{-1}).$$

High-order Yule-Walker (HOYW)

- Start from ARMA-model:

$$A(z)y(t) = A(z)e(t), \quad A(z) = \prod_{k=1}^{n}(1 - e^{i\omega_k} z^{-1}).$$

- Can't divide away $A(z)$ due to poles on unit circle!

High-order Yule-Walker (HOYW)

- Start from ARMA-model:

$$A(z)y(t) = A(z)e(t), \quad A(z) = \prod_{k=1}^{n}(1 - e^{i\omega_k} z^{-1}).$$

- Can't divide away $A(z)$ due to poles on unit circle!
- Multiply both sides with some filter $\tilde{A}(z)$ to get

$$B(z)y(t) = B(z)e(t),$$

where $B(z) = A(z)\tilde{A}(z)$ has order $L > n$ ("high order").

High-order Yule-Walker (HOYW)

- Start from ARMA-model:

$$A(z)y(t) = A(z)e(t), \quad A(z) = \prod_{k=1}^{n}(1 - e^{i\omega_k} z^{-1}).$$

- Can't divide away $A(z)$ due to poles on unit circle!
- Multiply both sides with some filter $\tilde{A}(z)$ to get

$$B(z)y(t) = B(z)e(t),$$

  where $B(z) = A(z)\tilde{A}(z)$ has order $L > n$ ("high order").
- $B(z)$ will have n roots $\{e^{i\omega_k}\}_{k=1}^{n}$ on the unit circle corresponding to the frequencies, and $L - n$ spurious roots.

UPPSALA UNIVERSITET

High-order Yule-Walker (HOYW)

- Start from ARMA-model:

$$A(z)y(t) = A(z)e(t), \quad A(z) = \prod_{k=1}^{n}(1 - e^{i\omega_k} z^{-1}).$$

- Can't divide away $A(z)$ due to poles on unit circle!
- Multiply both sides with some filter $\tilde{A}(z)$ to get

$$B(z)y(t) = B(z)e(t),$$

where $B(z) = A(z)\tilde{A}(z)$ has order $L > n$ ("high order").
- $B(z)$ will have n roots $\{e^{i\omega_k}\}_{k=1}^{n}$ on the unit circle corresponding to the frequencies, and $L - n$ spurious roots.
- Idea: Estimate $B(z)$ and take the n roots $\{r_k\}_{k=1}^{n}$ closest to the unit circle to get estimates $\hat{\omega}_k = \arg(r_k)$.

per.mattssonk@it.uu.se

High-order YW-equations

$$B(z)y(t) = B(z)e(t).$$

- From before we know that

$$r(k) + \sum_{i=1}^{L} b_i r(k-i) = 0, k > L.$$

UPPSALA UNIVERSITET

High-order YW-equations

$$B(z)y(t) = B(z)e(t).$$

- From before we know that

$$r(k) + \sum_{i=1}^{L} b_i r(k-i) = 0, k > L.$$

- For $k = L+1, \dots, L+M$ we get

$$\underbrace{\begin{bmatrix} r(L) & \cdots & r(1) \\ \vdots & \ddots & \vdots \\ r(L+M-1) & \cdots & r(M) \end{bmatrix}}_{\Omega} b = -\underbrace{\begin{bmatrix} r(L+1) \\ \vdots \\ r(L+M) \end{bmatrix}}_{\rho}$$

per.mattssonk@it.uu.se

High-order YW-equations

$$B(z)y(t) = B(z)e(t).$$

- From before we know that

$$r(k) + \sum_{i=1}^{L} b_i r(k-i) = 0, k > L.$$

- For $k = L+1, \ldots, L+M$ we get

$$\underbrace{\begin{bmatrix} r(L) & \cdots & r(1) \\ \vdots & \ddots & \vdots \\ r(L+M-1) & \cdots & r(M) \end{bmatrix}}_{\Omega} b = -\underbrace{\begin{bmatrix} r(L+1) \\ \vdots \\ r(L+M) \end{bmatrix}}_{\rho}$$

- Important property: $\text{rank}(\Omega) = n$.

per.mattssonk@it.uu.se

High-order YW-equations

$$B(z)y(t) = B(z)e(t).$$

- From before we know that
$$r(k) + \sum_{i=1}^{L} b_i r(k-i) = 0, k > L.$$
- For $k = L+1, \ldots, L+M$ we get
$$\underbrace{\begin{bmatrix} r(L) & \cdots & r(1) \\ \vdots & \ddots & \vdots \\ r(L+M-1) & \cdots & r(M) \end{bmatrix}}_{\Omega} b = -\underbrace{\begin{bmatrix} r(L+1) \\ \vdots \\ r(L+M) \end{bmatrix}}_{\rho}$$
- Important property: $\text{rank}(\Omega) = n$.
- When $L > n$ this rank-condition gives extra information about the structure of Ω.

HOYW-equations with estimated ACS

- Form $\widehat{\Omega}$ and $\hat{\rho}$ by replacing $r(k)$ with $\hat{r}(k)$ in Ω and ρ,

$$\widehat{\Omega} b \approx -\hat{\rho}. \tag{1}$$

HOYW-equations with estimated ACS

- Form $\widehat{\Omega}$ and $\hat{\rho}$ by replacing $r(k)$ with $\hat{r}(k)$ in Ω and ρ,
$$\widehat{\Omega} b \approx -\hat{\rho}. \tag{1}$$
- Random errors in $\hat{r}(k)$, so we (almost surely) get a full-rank $\widehat{\Omega}$,
$$\text{rank}(\widehat{\Omega}) = \min(M, L) > n.$$

HOYW-equations with estimated ACS

- Form $\widehat{\Omega}$ and $\hat{\rho}$ by replacing $r(k)$ with $\hat{r}(k)$ in Ω and ρ,
$$\widehat{\Omega} b \approx -\hat{\rho}. \tag{1}$$
- Random errors in $\hat{r}(k)$, so we (almost surely) get a full-rank $\widehat{\Omega}$,
$$\text{rank}(\widehat{\Omega}) = \min(M, L) > n.$$
- Ill-conditioned: For reasonable large N, $\widehat{\Omega}$ will be close to a rank-n matrix. Hence, solving for b in (1) directly is numerically problematic.

HOYW-equations with estimated ACS

- Form $\widehat{\Omega}$ and $\hat{\rho}$ by replacing $r(k)$ with $\hat{r}(k)$ in Ω and ρ,
$$\widehat{\Omega} b \approx -\hat{\rho}. \tag{1}$$
- Random errors in $\hat{r}(k)$, so we (almost surely) get a full-rank $\widehat{\Omega}$,
$$\text{rank}(\widehat{\Omega}) = \min(M, L) > n.$$
- Ill-conditioned: For reasonable large N, $\widehat{\Omega}$ will be close to a rank-n matrix. Hence, solving for b in (1) directly is numerically problematic.
- Idea: First approximate $\widehat{\Omega}$ with a rank-n matrix $\widehat{\Omega}_n$.

HOYW-equations with estimated ACS

- Form $\widehat{\Omega}$ and $\hat{\rho}$ by replacing $r(k)$ with $\hat{r}(k)$ in Ω and ρ,
$$\widehat{\Omega} b \approx -\hat{\rho}. \tag{1}$$
- Random errors in $\hat{r}(k)$, so we (almost surely) get a full-rank $\widehat{\Omega}$,
$$\text{rank}(\widehat{\Omega}) = \min(M, L) > n.$$
- Ill-conditioned: For reasonable large N, $\widehat{\Omega}$ will be close to a rank-n matrix. Hence, solving for b in (1) directly is numerically problematic.
- Idea: First approximate $\widehat{\Omega}$ with a rank-n matrix $\widehat{\Omega}_n$.
- $\widehat{\Omega}_n$ should be closer to the true Ω since we exploited
$$\text{rank}(\Omega) = n.$$

per.mattssonk@it.uu.se

UPPSALA UNIVERSITET

HOYW-equations with estimated ACS

- Form $\widehat{\Omega}$ and $\hat{\rho}$ by replacing $r(k)$ with $\hat{r}(k)$ in Ω and ρ,
$$\widehat{\Omega} b \approx -\hat{\rho}. \tag{1}$$
- Random errors in $\hat{r}(k)$, so we (almost surely) get a full-rank $\widehat{\Omega}$,
$$\text{rank}(\widehat{\Omega}) = \min(M, L) > n.$$
- Ill-conditioned: For reasonable large N, $\widehat{\Omega}$ will be close to a rank-n matrix. Hence, solving for b in (1) directly is numerically problematic.
- Idea: First approximate $\widehat{\Omega}$ with a rank-n matrix $\widehat{\Omega}_n$.
- $\widehat{\Omega}_n$ should be closer to the true Ω since we exploited
$$\text{rank}(\Omega) = n.$$
- Find $\hat{b}$ by solving (LS)
$$\widehat{\Omega}_n \hat{b} \approx -\hat{\rho} \Longrightarrow \hat{b} = -\widehat{\Omega}_n^{\dagger} \hat{\rho}.$$

per.mattssonk@it.uu.se

HOYW in practice

- Use SVD on $\widehat{\Omega}$,

$$\widehat{\Omega} = U\Sigma V^* = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} V_1^* \\ V_2^* \end{bmatrix},$$

where Σ_1 contains the n largest singular values.

HOYW in practice

- Use SVD on $\widehat{\Omega}$,

$$\widehat{\Omega} = U\Sigma V^* = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} V_1^* \\ V_2^* \end{bmatrix},$$

  where Σ_1 contains the n largest singular values.
- Truncate (i.e. replace Σ_2 with 0) to get rank-n matrix

$$\widehat{\Omega}_n = U_1 \Sigma_1 V_1^* \approx \widehat{\Omega}.$$

HOYW in practice

- Use SVD on $\widehat{\Omega}$,

$$\widehat{\Omega} = U\Sigma V^* = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} V_1^* \\ V_2^* \end{bmatrix},$$

where Σ_1 contains the n largest singular values.

- Truncate (i.e. replace Σ_2 with 0) to get rank-n matrix

$$\widehat{\Omega}_n = U_1 \Sigma_1 V_1^* \approx \widehat{\Omega}.$$

- Solve for $\hat{b}$,

$$\hat{b} = -\widehat{\Omega}_n^{\dagger} \hat{\rho} =$$

HOYW in practice

- Use SVD on $\widehat{\Omega}$,

$$\widehat{\Omega} = U\Sigma V^* = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} V_1^* \\ V_2^* \end{bmatrix},$$

  where Σ_1 contains the n largest singular values.
- Truncate (i.e. replace Σ_2 with 0) to get rank-n matrix

$$\widehat{\Omega}_n = U_1 \Sigma_1 V_1^* \approx \widehat{\Omega}.$$

- Solve for $\hat{b}$,

$$\hat{b} = -\widehat{\Omega}_n^{\dagger} \hat{\rho} = -V_1 \Sigma_1^{-1} U_1^* \hat{\rho}.$$

HOYW in practice

- Use SVD on $\widehat{\Omega}$,

$$\widehat{\Omega} = U\Sigma V^* = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} V_1^* \\ V_2^* \end{bmatrix},$$

  where Σ_1 contains the n largest singular values.
- Truncate (i.e. replace Σ_2 with 0) to get rank-n matrix

$$\widehat{\Omega}_n = U_1 \Sigma_1 V_1^* \approx \widehat{\Omega}.$$

- Solve for $\hat{b}$,

$$\hat{b} = -\widehat{\Omega}_n^{\dagger} \hat{\rho} = -V_1 \Sigma_1^{-1} U_1^* \hat{\rho}.$$

- $\{\hat{\omega}_k\}_{k=1}^{n}$ from the n roots of $\hat{B}(z)$ closest to the unit circle.

HOYW in practice

- Use SVD on $\widehat{\Omega}$,

$$\widehat{\Omega} = U\Sigma V^* = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} V_1^* \\ V_2^* \end{bmatrix},$$

  where Σ_1 contains the n largest singular values.

- Truncate (i.e. replace Σ_2 with 0) to get rank-n matrix

$$\widehat{\Omega}_n = U_1 \Sigma_1 V_1^* \approx \widehat{\Omega}.$$

- Solve for $\hat{b}$,

$$\hat{b} = -\widehat{\Omega}_n^{\dagger} \hat{\rho} = -V_1 \Sigma_1^{-1} U_1^* \hat{\rho}.$$

- $\{\hat{\omega}_k\}_{k=1}^{n}$ from the n roots of $\hat{B}(z)$ closest to the unit circle.
- User parameters: $M \approx L$, and $L + M$ a fraction of N.

per.mattssonk@it.uu.se

HOYW in practice

- Use SVD on $\widehat{\Omega}$,

$$\widehat{\Omega} = U\Sigma V^* = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} V_1^* \\ V_2^* \end{bmatrix},$$

  where Σ_1 contains the n largest singular values.
- Truncate (i.e. replace Σ_2 with 0) to get rank-n matrix

$$\widehat{\Omega}_n = U_1 \Sigma_1 V_1^* \approx \widehat{\Omega}.$$

- Solve for $\hat{b}$,

$$\hat{b} = -\widehat{\Omega}_n^\dagger \hat{\rho} = -V_1 \Sigma_1^{-1} U_1^* \hat{\rho}.$$

- $\{\hat{\omega}_k\}_{k=1}^n$ from the n roots of $\hat{B}(z)$ closest to the unit circle.
- User parameters: $M \approx L$, and $L + M$ a fraction of N.
- Note: SVD can also give an idea about the size of n.

MUltiple SIgnal Classification (MUSIC)

The Covariance Matrix Equation

Let

$$a(\omega) = \begin{bmatrix} 1 & e^{-i\omega} & \cdots & e^{-i(m-1)\omega} \end{bmatrix}^\top \text{ for some } m > n$$
$$A = \begin{bmatrix} a(\omega_1) & \cdots & a(\omega_n) \end{bmatrix} \in \mathbb{C}^{m \times n}$$

Let $\tilde{y}(t) = \begin{bmatrix} y(t) & \cdots & y(t-m+1) \end{bmatrix}^\top$, then

$$\tilde{y}(t) = A\tilde{x}(t) + \tilde{e}(t)$$
$$\tilde{x}(t) = \begin{bmatrix} x_1(t) & \cdots & x_n(t) \end{bmatrix}^\top, \quad \tilde{e}(t) = \begin{bmatrix} e(t) & \cdots & e(t-m+1) \end{bmatrix}^\top$$

Covariance matrix: $R = E\{\tilde{y}(t)\tilde{y}^*(t)\} = APA^* + \sigma^2 I$, where

$$P = E\{\tilde{x}(t)\tilde{x}^*(t)\} = \begin{bmatrix} \alpha_1^2 & & 0 \\ & \ddots & \\ 0 & & \alpha_n^2 \end{bmatrix}$$

per.mattssonk@it.uu.se

Some rank conditions

- $\text{rank}(A) = n$ (if $m \geq n$ and $\omega_k \neq \omega_p$ for $k \neq p$).

UPPSALA UNIVERSITET

Some rank conditions

- $\text{rank}(A) = n$ (if $m \geq n$ and $\omega_k \neq \omega_p$ for $k \neq p$).
- Hence,

$$\text{rank}(APA^*) = n.$$

per.mattssonk@it.uu.se

UPPSALA UNIVERSITET

Some rank conditions

- $\text{rank}(A) = n$ (if $m \geq n$ and $\omega_k \neq \omega_p$ for $k \neq p$).
- Hence,
$$\text{rank}(APA^*) = n.$$
- APA^* has n strictly positive eigenvalues, and $m - n$ eigenvalues equal to 0,

Some rank conditions

- $\text{rank}(A) = n$ (if $m \geq n$ and $\omega_k \neq \omega_p$ for $k \neq p$).
- Hence,
$$\text{rank}(APA^*) = n.$$
- APA^* has n strictly positive eigenvalues, and $m - n$ eigenvalues equal to 0,
$$\tilde{\lambda}_1 \geq \tilde{\lambda}_2 \geq \cdots \geq \tilde{\lambda}_n > 0, \quad \tilde{\lambda}_{n+1} = \cdots = \tilde{\lambda}_m = 0.$$

Some rank conditions

- $\text{rank}(A) = n$ (if $m \geq n$ and $\omega_k \neq \omega_p$ for $k \neq p$).
- Hence,
$$\text{rank}(APA^*) = n.$$
- APA^* has n strictly positive eigenvalues, and $m - n$ eigenvalues equal to 0,
$$\tilde{\lambda}_1 \geq \tilde{\lambda}_2 \geq \cdots \geq \tilde{\lambda}_n > 0, \quad \tilde{\lambda}_{n+1} = \cdots = \tilde{\lambda}_m = 0.$$
- The eigenvalues λ_k of $R = APA^* + \sigma^2 I$ are thus given by
$$\lambda_k =$$

per.mattssonk@it.uu.se

Some rank conditions

- rank$(A) = n$ (if $m \geq n$ and $\omega_k \neq \omega_p$ for $k \neq p$).
- Hence,
$$\text{rank}(APA^*) = n.$$
- APA^* has n strictly positive eigenvalues, and $m - n$ eigenvalues equal to 0,
$$\tilde{\lambda}_1 \geq \tilde{\lambda}_2 \geq \cdots \geq \tilde{\lambda}_n > 0, \quad \tilde{\lambda}_{n+1} = \cdots = \tilde{\lambda}_m = 0.$$
- The eigenvalues λ_k of $R = APA^* + \sigma^2 I$ are thus given by
$$\lambda_k = \tilde{\lambda}_k + \sigma^2 > 0, \quad k = 1, \ldots, m,$$

per.mattssonk@it.uu.se

Some rank conditions

- rank$(A) = n$ (if $m \geq n$ and $\omega_k \neq \omega_p$ for $k \neq p$).
- Hence,
$$\text{rank}(APA^*) = n.$$
- APA^* has n strictly positive eigenvalues, and $m - n$ eigenvalues equal to 0,
$$\tilde{\lambda}_1 \geq \tilde{\lambda}_2 \geq \cdots \geq \tilde{\lambda}_n > 0, \quad \tilde{\lambda}_{n+1} = \cdots = \tilde{\lambda}_m = 0.$$
- The eigenvalues λ_k of $R = APA^* + \sigma^2 I$ are thus given by
$$\lambda_k = \tilde{\lambda}_k + \sigma^2 > 0, \quad k = 1, \ldots, m,$$
- and thus
$$\text{rank}(R) = m.$$

Eigendecomposition of R

$$R = APA^* + \sigma^2 I \quad (m > n)$$

- Eigenvalues: $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n > \lambda_{n+1} = \cdots = \lambda_m = \sigma^2$

UPPSALA UNIVERSITET

Eigendecomposition of R

$$R = APA^* + \sigma^2 I \quad (m > n)$$

- Eigenvalues: $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n > \lambda_{n+1} = \cdots = \lambda_m = \sigma^2$
- Eigenvectors: $s_1, \ldots, s_n,\ g_1, \ldots, g_{m-n}$ orthonormal

$$S = \begin{bmatrix} s_1 & \cdots & s_n \end{bmatrix} \text{ corresponding to } \lambda_1, \ldots, \lambda_n$$

$$G = \begin{bmatrix} g_1 & \cdots & g_{m-n} \end{bmatrix} \text{ corresponding to } \lambda_{n+1}, \ldots, \lambda_m$$

per.mattssonk@it.uu.se

Eigendecomposition of R

$$R = APA^* + \sigma^2 I \quad (m > n)$$

- Eigenvalues: $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n > \lambda_{n+1} = \cdots = \lambda_m = \sigma^2$
- Eigenvectors: $s_1, \ldots, s_n,\ g_1, \ldots, g_{m-n}$ orthonormal

$$S = \begin{bmatrix} s_1 & \cdots & s_n \end{bmatrix} \text{ corresponding to } \lambda_1, \ldots, \lambda_n$$

$$G = \begin{bmatrix} g_1 & \cdots & g_{m-n} \end{bmatrix} \text{ corresponding to } \lambda_{n+1}, \ldots, \lambda_m$$

- Eigendecomposition:

$$R = \begin{bmatrix} S & G \end{bmatrix} \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_m \end{bmatrix} \begin{bmatrix} S^* \\ G^* \end{bmatrix}.$$

MUSIC, motivation

- Note that

$$RG =$$

MUSIC, motivation

- Note that

$$RG = \sigma^2 G =$$

per.mattssonk@it.uu.se

MUSIC, motivation

- Note that
$$RG = \sigma^2 G = APA^* G + \sigma^2 G$$

per.mattssonk@it.uu.se

MUSIC, motivation

- Note that
$$RG = \sigma^2 G = APA^* G + \sigma^2 G$$
- Follows that $APA^* G = 0$.

MUSIC, motivation

- Note that
$$RG = \sigma^2 G = APA^* G + \sigma^2 G$$

- Follows that $APA^* G = 0$. Since AP has full rank we get

$$0 = A^* G =$$

MUSIC, motivation

- Note that
$$RG = \sigma^2 G = APA^*G + \sigma^2 G$$

- Follows that $APA^*G = 0$. Since AP has full rank we get
$$0 = A^*G = \begin{bmatrix} a^*(\omega_1)G \\ \vdots \\ a^*(\omega_n)G \end{bmatrix}$$

UPPSALA UNIVERSITET

MUSIC, motivation

- Note that
$$RG = \sigma^2 G = APA^*G + \sigma^2 G$$

- Follows that $APA^*G = 0$. Since AP has full rank we get
$$0 = A^*G = \begin{bmatrix} a^*(\omega_1)G \\ \vdots \\ a^*(\omega_n)G \end{bmatrix} \implies a^*(\omega)G = 0 \text{ if } \omega = \omega_k.$$

per.mattssonk@it.uu.se

UPPSALA UNIVERSITET

MUSIC, motivation

- Note that
$$RG = \sigma^2 G = APA^*G + \sigma^2 G$$
- Follows that $APA^*G = 0$. Since AP has full rank we get
$$0 = A^*G = \begin{bmatrix} a^*(\omega_1)G \\ \vdots \\ a^*(\omega_n)G \end{bmatrix} \implies a^*(\omega)G = 0 \text{ if } \omega = \omega_k.$$
- Result: For $\omega = \omega_k$ (the true frequencies)
$$a^*(\omega)GG^*a(\omega) = 0$$

UPPSALA UNIVERSITET

MUSIC, motivation

- Note that
$$RG = \sigma^2 G = APA^*G + \sigma^2 G$$
- Follows that $APA^*G = 0$. Since AP has full rank we get
$$0 = A^*G = \begin{bmatrix} a^*(\omega_1)G \\ \vdots \\ a^*(\omega_n)G \end{bmatrix} \implies a^*(\omega)G = 0 \text{ if } \omega = \omega_k.$$
- Result: For $\omega = \omega_k$ (the true frequencies)
$$a^*(\omega)GG^*a(\omega) = 0$$
- It can be shown that the above holds **if and only if** $\omega = \omega_k$ for $k = 1, \ldots, n$.

MUSIC, motivation

- Note that
$$RG = \sigma^2 G = APA^* G + \sigma^2 G$$
- Follows that $APA^* G = 0$. Since AP has full rank we get
$$0 = A^* G = \begin{bmatrix} a^*(\omega_1) G \\ \vdots \\ a^*(\omega_n) G \end{bmatrix} \implies a^*(\omega) G = 0 \text{ if } \omega = \omega_k.$$
- Result: For $\omega = \omega_k$ (the true frequencies)
$$a^*(\omega) G G^* a(\omega) = 0$$
- It can be shown that the above holds **if and only if** $\omega = \omega_k$ for $k = 1, \dots, n$.
- (Any $a(\tilde{\omega})$ that satisfy the equation belongs to $\mathcal{N}(G^*)$, which is a subspace of dimension n. At most n vectors on the form $a(\omega)$ can belong to this subspace, which by the above must be $\{a(\omega_k)\}_{k=1}^n$.)

MUSIC

Compute sample covariance matrix (hermitian positive definite)

$$\hat{R} = \frac{1}{N} \sum_{t=m}^{N} \tilde{y}(t)\tilde{y}^*(t)$$

and its eigendecomposition (here SVD) giving matrices $\hat{S}$ and $\hat{G}$

MUSIC

Compute sample covariance matrix (hermitian positive definite)

$$\hat{R} = \frac{1}{N} \sum_{t=m}^{N} \tilde{y}(t)\tilde{y}^*(t)$$

and its eigendecomposition (here SVD) giving matrices $\hat{S}$ and $\hat{G}$

Spectral MUSIC

Plot pseudo spectrum and determine frequencies from the n highest peaks

$$\frac{1}{a^*(\omega)\hat{G}\hat{G}^*a(\omega)} \quad \omega \in [-\pi, \pi]$$

MUSIC

Compute sample covariance matrix (hermitian positive definite)

$$\hat{R} = \frac{1}{N} \sum_{t=m}^{N} \tilde{y}(t)\tilde{y}^*(t)$$

and its eigendecomposition (here SVD) giving matrices $\hat{S}$ and $\hat{G}$

Spectral MUSIC

Plot pseudo spectrum and determine frequencies from the n highest peaks

$$\frac{1}{a^*(\omega)\hat{G}\hat{G}^* a(\omega)} \quad \omega \in [-\pi, \pi]$$

Root MUSIC

Determine frequencies from the angular positions of the n roots of

$$a^{\mathsf{T}}(z^{-1})\hat{G}\hat{G}^* a(z) = 0$$

closest to the unit circle

Min-Norm method

UPPSALA UNIVERSITET

The Min-Norm method

- Goal: Reduce computational burden, and reduce risk of false frequency estimates.

The Min-Norm method

- Goal: Reduce computational burden, and reduce risk of false frequency estimates.
- Idea: From our previous discussion we can see that for any vector $g \in \mathcal{R}(G) = \mathcal{N}(A^*)$ we have

$$a^*(\omega)g = 0, \quad \text{for } \omega = \omega_k.$$

The Min-Norm method

- Goal: Reduce computational burden, and reduce risk of false frequency estimates.
- Idea: From our previous discussion we can see that for any vector $g \in \mathcal{R}(G) = \mathcal{N}(A^*)$ we have

$$a^*(\omega)g = 0, \quad \text{for } \omega = \omega_k.$$

- Computationally easier to solve this equation!

per.mattssonk@it.uu.se

UPPSALA UNIVERSITET

The Min-Norm method

- Goal: Reduce computational burden, and reduce risk of false frequency estimates.
- Idea: From our previous discussion we can see that for any vector $g \in \mathcal{R}(G) = \mathcal{N}(A^*)$ we have

$$a^*(\omega)g = 0, \quad \text{for } \omega = \omega_k.$$

- Computationally easier to solve this equation!
- Which vector to use? Are strong empirical evidence that the vector

$$g = \begin{bmatrix} 1 \\ \tilde{g} \end{bmatrix} \in \mathcal{R}(G)$$

with minimum Euclidean norm have good properties (accuracy, and reduced risk of spurious ω-estimates).

Min-Norm

Find

$$\begin{bmatrix} 1 \\ \hat{g} \end{bmatrix} \in \mathcal{R}(\hat{G})$$

with minimum Euclidean norm (see textbook).

UPPSALA UNIVERSITET

Min-Norm

Find

$$\begin{bmatrix} 1 \\ \hat{g} \end{bmatrix} \in \mathcal{R}(\hat{G})$$

with minimum Euclidean norm (see textbook).

Spectral Min-Norm

Plot pseudo spectrum and determine frequencies from the n highest peaks

$$\frac{1}{a^*(\omega) \begin{bmatrix} 1 \\ \hat{g} \end{bmatrix}}$$

per.mattssonk@it.uu.se

Min-Norm

Find

$$\begin{bmatrix} 1 \\ \hat{g} \end{bmatrix} \in \mathcal{R}(\hat{G})$$

with minimum Euclidean norm (see textbook).

Spectral Min-Norm

Plot pseudo spectrum and determine frequencies from the n highest peaks

$$\frac{1}{a^*(\omega) \begin{bmatrix} 1 \\ \hat{g} \end{bmatrix}}$$

Root Min-Norm

Determine frequencies from the angular positions of the n roots of

$$a^{\mathsf{T}}(z^{-1}) \begin{bmatrix} 1 \\ \hat{g} \end{bmatrix} = 0$$

closest to the unit circle

Estimation of Signal Parameters by Rotational Invariance Techniques (ESPRIT)

esprit (noun) – vivacious cleverness or wit

Theory of ESPRIT

- Let A_1 be A with last row removed, and A_2 be A with first row removed

Theory of ESPRIT

- Let A_1 be A with last row removed, and A_2 be A with first row removed, i.e.,

$$A_1 = \begin{bmatrix} I_{m-1} & 0 \end{bmatrix} A, \quad A_2 = \begin{bmatrix} 0 & I_{m-1} \end{bmatrix} A, \quad (m-1 \times n)$$

Theory of ESPRIT

- Let A_1 be A with last row removed, and A_2 be A with first row removed, i.e.,

$$A_1 = \begin{bmatrix} I_{m-1} & 0 \end{bmatrix} A, \quad A_2 = \begin{bmatrix} 0 & I_{m-1} \end{bmatrix} A, \quad (m-1 \times n)$$

- then,

$$A_2 = A_1 D, \quad D = \begin{bmatrix} e^{-i\omega_1} & & 0 \\ & \ddots & \\ 0 & & e^{-i\omega_n} \end{bmatrix}.$$

Theory of ESPRIT

- Let A_1 be A with last row removed, and A_2 be A with first row removed, i.e.,

$$A_1 = \begin{bmatrix} I_{m-1} & 0 \end{bmatrix} A, \quad A_2 = \begin{bmatrix} 0 & I_{m-1} \end{bmatrix} A, \quad (m - 1 \times n)$$

- then,

$$A_2 = A_1 D, \quad D = \begin{bmatrix} e^{-i\omega_1} & & 0 \\ & \ddots & \\ 0 & & e^{-i\omega_n} \end{bmatrix}.$$

- In the same way let

$$S_1 = \begin{bmatrix} I_{m-1} & 0 \end{bmatrix} S, \quad S_2 = \begin{bmatrix} 0 & I_{m-1} \end{bmatrix} S.$$

- There exist nonsingular C such that $S = AC$.

UPPSALA UNIVERSITET

Theory of ESPRIT

- Let A_1 be A with last row removed, and A_2 be A with first row removed, i.e.,

$$A_1 = \begin{bmatrix} I_{m-1} & 0 \end{bmatrix} A, \quad A_2 = \begin{bmatrix} 0 & I_{m-1} \end{bmatrix} A, \quad (m-1 \times n)$$

- then,

$$A_2 = A_1 D, \quad D = \begin{bmatrix} e^{-i\omega_1} & & 0 \\ & \ddots & \\ 0 & & e^{-i\omega_n} \end{bmatrix}.$$

- In the same way let

$$S_1 = \begin{bmatrix} I_{m-1} & 0 \end{bmatrix} S, \quad S_2 = \begin{bmatrix} 0 & I_{m-1} \end{bmatrix} S.$$

- There exist nonsingular C such that $S = AC$. Hence,

$$S_2 =$$

per.mattssonk@it.uu.se

Theory of ESPRIT

- Let A_1 be A with last row removed, and A_2 be A with first row removed, i.e.,

$$A_1 = \begin{bmatrix} I_{m-1} & 0 \end{bmatrix} A, \quad A_2 = \begin{bmatrix} 0 & I_{m-1} \end{bmatrix} A, \quad (m-1 \times n)$$

- then,

$$A_2 = A_1 D, \quad D = \begin{bmatrix} e^{-i\omega_1} & & 0 \\ & \ddots & \\ 0 & & e^{-i\omega_n} \end{bmatrix}.$$

- In the same way let

$$S_1 = \begin{bmatrix} I_{m-1} & 0 \end{bmatrix} S, \quad S_2 = \begin{bmatrix} 0 & I_{m-1} \end{bmatrix} S.$$

- There exist nonsingular C such that $S = AC$. Hence,

$$S_2 = A_2 C =$$

Theory of ESPRIT

- Let A_1 be A with last row removed, and A_2 be A with first row removed, i.e.,

$$A_1 = \begin{bmatrix} I_{m-1} & 0 \end{bmatrix} A, \quad A_2 = \begin{bmatrix} 0 & I_{m-1} \end{bmatrix} A, \quad (m-1 \times n)$$

- then,

$$A_2 = A_1 D, \quad D = \begin{bmatrix} e^{-i\omega_1} & & 0 \\ & \ddots & \\ 0 & & e^{-i\omega_n} \end{bmatrix}.$$

- In the same way let

$$S_1 = \begin{bmatrix} I_{m-1} & 0 \end{bmatrix} S, \quad S_2 = \begin{bmatrix} 0 & I_{m-1} \end{bmatrix} S.$$

- There exist nonsingular C such that $S = AC$. Hence,

$$S_2 = A_2 C = A_1 D C =$$

Theory of ESPRIT

- Let A_1 be A with last row removed, and A_2 be A with first row removed, i.e.,

$$A_1 = \begin{bmatrix} I_{m-1} & 0 \end{bmatrix} A, \quad A_2 = \begin{bmatrix} 0 & I_{m-1} \end{bmatrix} A, \quad (m-1 \times n)$$

- then,

$$A_2 = A_1 D, \quad D = \begin{bmatrix} e^{-i\omega_1} & & 0 \\ & \ddots & \\ 0 & & e^{-i\omega_n} \end{bmatrix}.$$

- In the same way let

$$S_1 = \begin{bmatrix} I_{m-1} & 0 \end{bmatrix} S, \quad S_2 = \begin{bmatrix} 0 & I_{m-1} \end{bmatrix} S.$$

- There exist nonsingular C such that $S = AC$. Hence,

$$S_2 = A_2 C = A_1 D C = S_1 \underbrace{C^{-1} D C}_{\phi}.$$

ESPIRIT

- We have
$$S_2 = S_1\phi, \quad \text{where } \phi = C^{-1}DC.$$

ESPIRIT

- We have
$$S_2 = S_1\phi, \quad \text{where } \phi = C^{-1}DC.$$
- Eigenvalues of ϕ: ϕ is similar to $D \Rightarrow$

ESPIRIT

- We have
$$S_2 = S_1\phi, \quad \text{where } \phi = C^{-1}DC.$$
- Eigenvalues of ϕ: ϕ is similar to D $\Rightarrow$The eigenvalues of ϕ are the same as the eigenvalues of D, i.e., $\{e^{-i\omega_k}\}_{k=1}^{n}$.

ESPIRIT

- We have
$$S_2 = S_1\phi, \quad \text{where } \phi = C^{-1}DC.$$
- Eigenvalues of ϕ: ϕ is similar to D $\Rightarrow$The eigenvalues of ϕ are the same as the eigenvalues of D, i.e., $\{e^{-i\omega_k}\}_{k=1}^{n}$.
- S_1 and S_2 have full rank, so ϕ is uniquely given by
$$\phi = (S_1^* S_1)^{-1} S_1^* S_2.$$

UPPSALA UNIVERSITET

ESPIRIT

- We have
$$S_2 = S_1\phi, \quad \text{where } \phi = C^{-1}DC.$$
- Eigenvalues of ϕ: ϕ is similar to D $\Rightarrow$The eigenvalues of ϕ are the same as the eigenvalues of D, i.e., $\{e^{-i\omega_k}\}_{k=1}^{n}$.
- S_1 and S_2 have full rank, so ϕ is uniquely given by
$$\phi = (S_1^* S_1)^{-1} S_1^* S_2.$$

ESPIRIT:

- Compute estimated $\hat{R}$ for some $m > n$.

ESPIRIT

- We have
$$S_2 = S_1\phi, \quad \text{where } \phi = C^{-1}DC.$$
- Eigenvalues of ϕ: ϕ is similar to D $\Rightarrow$The eigenvalues of ϕ are the same as the eigenvalues of D, i.e., $\{e^{-i\omega_k}\}_{k=1}^n$.
- S_1 and S_2 have full rank, so ϕ is uniquely given by
$$\phi = (S_1^* S_1)^{-1} S_1^* S_2.$$

ESPIRIT:

- Compute estimated $\hat{R}$ for some $m > n$.
- Use SVD to get $\hat{S}$ and construct $\hat{S}_1$ and $\hat{S}_2$.

ESPIRIT

- We have
$$S_2 = S_1\phi, \quad \text{where } \phi = C^{-1}DC.$$
- Eigenvalues of ϕ: ϕ is similar to D $\Rightarrow$The eigenvalues of ϕ are the same as the eigenvalues of D, i.e., $\{e^{-i\omega_k}\}_{k=1}^{n}$.
- S_1 and S_2 have full rank, so ϕ is uniquely given by
$$\phi = (S_1^* S_1)^{-1} S_1^* S_2.$$

ESPIRIT:

- Compute estimated $\hat{R}$ for some $m > n$.
- Use SVD to get $\hat{S}$ and construct $\hat{S}_1$ and $\hat{S}_2$.
- Let $\hat{\phi} = (\hat{S}_1^* \hat{S}_1)^{-1} \hat{S}_1^* \hat{S}_2$ (LS-solution).

ESPIRIT

- We have
$$S_2 = S_1\phi, \quad \text{where } \phi = C^{-1}DC.$$
- Eigenvalues of ϕ: ϕ is similar to D $\Rightarrow$The eigenvalues of ϕ are the same as the eigenvalues of D, i.e., $\{e^{-i\omega_k}\}_{k=1}^{n}$.
- S_1 and S_2 have full rank, so ϕ is uniquely given by
$$\phi = (S_1^* S_1)^{-1} S_1^* S_2.$$

ESPIRIT:

- Compute estimated $\hat{R}$ for some $m > n$.
- Use SVD to get $\hat{S}$ and construct $\hat{S}_1$ and $\hat{S}_2$.
- Let $\hat{\phi} = (\hat{S}_1^* \hat{S}_1)^{-1} \hat{S}_1^* \hat{S}_2$ (LS-solution).
- Let $\hat{v}_k$ be the eigenvalues of $\hat{\phi}$. Then the estimated frequencies are
$$\hat{\omega}_k = -\arg(\hat{v}_k).$$

per.mattssonk@it.uu.se

UPPSALA UNIVERSITET

ESPIRIT

- We have
$$S_2 = S_1\phi, \quad \text{where } \phi = C^{-1}DC.$$
- Eigenvalues of ϕ: ϕ is similar to D $\Rightarrow$The eigenvalues of ϕ are the same as the eigenvalues of D, i.e., $\{e^{-i\omega_k}\}_{k=1}^{n}$.
- S_1 and S_2 have full rank, so ϕ is uniquely given by
$$\phi = (S_1^* S_1)^{-1} S_1^* S_2.$$

ESPIRIT:

- Compute estimated $\hat{R}$ for some $m > n$.
- Use SVD to get $\hat{S}$ and construct $\hat{S}_1$ and $\hat{S}_2$.
- Let $\hat{\phi} = (\hat{S}_1^* \hat{S}_1)^{-1} \hat{S}_1^* \hat{S}_2$ (LS-solution).
- Let $\hat{v}_k$ be the eigenvalues of $\hat{\phi}$. Then the estimated frequencies are
$$\hat{\omega}_k = -\arg(\hat{v}_k).$$
- If estimates of α_k are needed, use LS with the estimated frequencies.

Performance

- All methods can achieve super resolution (beyond the Periodogram)
- Finding the frequencies through subspace methods is tractable as opposed to NLS (typically)
- ESPRIT is typically slightly better than the others
- We still have to choose the order n

Method	Computational Burden	Accuracy / Resolution	Risk for False Freq Estimates
Periodogram	small	medium-high	medium
Nonlinear LS	very high	very high	very high
Yule-Walker	medium	high	medium
Pisarenko	small	low	none
MUSIC	high	high	medium
Min-Norm	medium	high	small
ESPRIT	medium	very high	none

MATLAB

Useful MATLAB functions:

- `hoyw()`
- `music()`
- `minnorm()`
- `esprit()`
- `freqaphi()`
- `lsa()`

Summary

- Alternatives to NLS, based on ARMA and covariance model
- HOYW, MUSIC, Min-Norm, ESPRIT
- Complicated derivation but easy to use (and implement)

Summary

- Alternatives to NLS, based on ARMA and covariance model
- HOYW, MUSIC, Min-Norm, ESPRIT
- Complicated derivation but easy to use (and implement)

Next lecture:

- How to choose the model order?
- Information criteria
- Heuristic approaches
- Cross validation