



دانشگاه تحصیلات تکمیلی صنعتی و فناوری پیشرفته
دانشکده مهندسی برق و کامپیوتر

تئوری تخمین

ESTIMATION THEORY

نیم سال اول سال تحصیلی ۰۴-۰۵

بهترین تخمینگر خطی بدون بایاس

2

□ مقدمه

□ تعریف BLUE

□ پیدا کردن BLUE

□ تعمیم به حالت برداری

- پیدا کردن تخمینگر MVU در بسیاری موارد ناممکن است، مثلاً وقتی که PDF داده ها را نمی دانیم.
- حتی اگر PDF داده ها را هم بدانیم، تضمینی برای یافتن تخمینگر MVU وجود ندارد.
- در چنین مواردی به سراغ تخمینگرهای «زیر بهینه» (sub-optimal) می رویم.
- یکی از این راه حل های زیربهینه، یافتن تخمینگری است که تابعی خطی از داده است و در ضمن بدون بایاس و داری کمترین واریانس است: **Best Linear Unbiased Estimator**
- برای یافتن BLUE تنها ممان اول و دوم داده کافی است و به PDF کامل نیاز نداریم.

□ در BLUE، تخمینگر تابعی خطی از داده است:

$$\hat{\theta} = \sum_{n=0}^{N-1} a_n x[n]$$

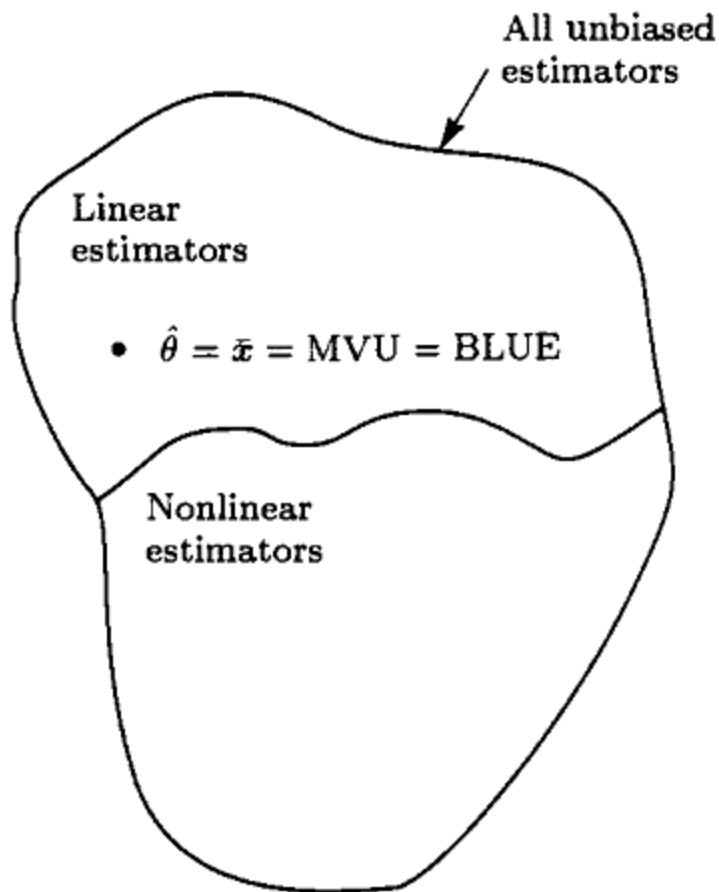
□ که a_n ها باید یافته شوند، به طوری که تخمینگر بدون بایاس و داری کمترین واریانس باشد.

□ اگر تخمینگر MVU تابعی خطی از داده باشد، آنگاه MVU و BLUE یکسان خواهند بود.

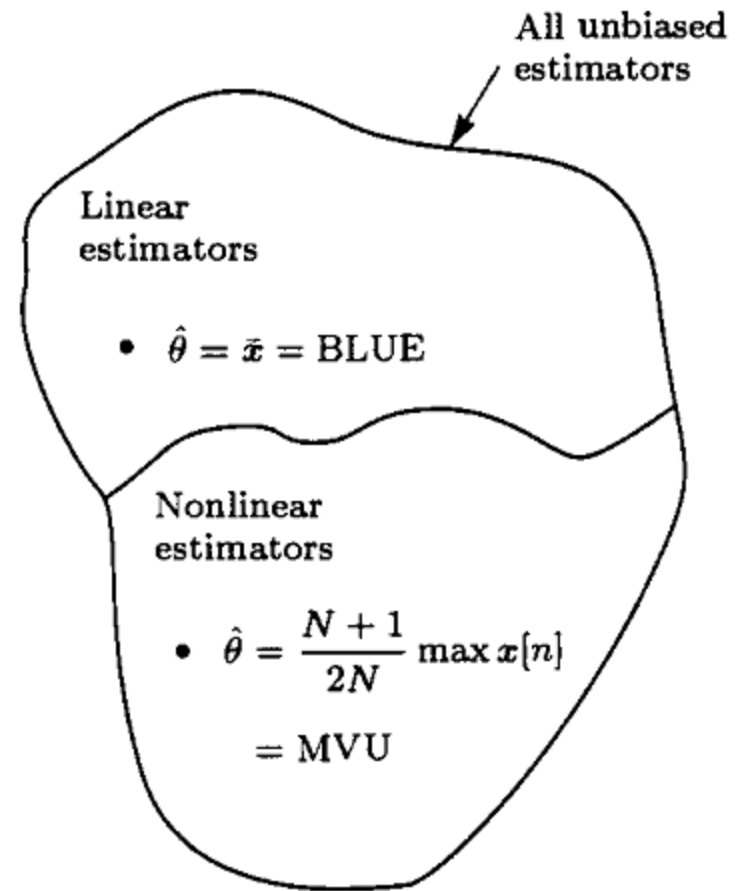
□ مثلاً در تخمین ولتاژ DC در نویز WGN تخمینگر MVU را به صورت زیر یافتیم:

$$\hat{\theta} = \bar{x} = \sum_{n=0}^{N-1} \frac{1}{N} x[n]$$

- مشخص است که این تخمینگر، تابعی خطی از داده است. بنابراین تخمینگر MVU و BLUE یکی هستند.
- ولی در مورد تخمین ولتاژ DC در نویز یکنواخت، تخمینگر MVU به صورت زیر به دست می آید:
$$\hat{\theta} = \frac{N+1}{2N} \max x[n]$$
- در اینجا مشخص است که MVU تابع خطی از داده نیست، در نتیجه BLUE یک راه حل زیربهبینه خواهد بود.
- با استفاده از حد کرامر-راو می توان فاصله واریانس BLUE را از بهترین واریانس ممکن در تخمینگرهای بدون بایاس به دست آورد.



(a) DC level in WGN; BLUE is optimal



(b) Mean of uniform noise; BLUE is suboptimal

□ همانطور که گفتیم در BLUE به دنبال ضرایبی هستیم که یک تخمینگر بدون بایاس و کمترین واریانس نتیجه دهد:

$$\hat{\theta} = \sum_{n=0}^{N-1} a_n x[n]$$

$$E(\hat{\theta}) = \sum_{n=0}^{N-1} a_n E(x[n]) = \theta.$$

□ شرط بدون بایاس بودن می گوید:

□ محاسبه واریانس تخمینگر:

$$\text{var}(\hat{\theta}) = E \left[\left(\sum_{n=0}^{N-1} a_n x[n] - E \left(\sum_{n=0}^{N-1} a_n x[n] \right) \right)^2 \right].$$

$$\begin{aligned} \text{var}(\hat{\theta}) &= E \left[(\mathbf{a}^T \mathbf{x} - \mathbf{a}^T E(\mathbf{x}))^2 \right] \\ &= E \left[(\mathbf{a}^T (\mathbf{x} - E(\mathbf{x})))^2 \right] \\ &= E \left[\mathbf{a}^T (\mathbf{x} - E(\mathbf{x})) (\mathbf{x} - E(\mathbf{x}))^T \mathbf{a} \right] \\ &= \mathbf{a}^T \mathbf{C} \mathbf{a}. \end{aligned}$$

□ حال باید بردار ضرایب $\mathbf{a}$ طوری پیدا شود که مقدار واریانس کمینه شود، به شرطی که:

$$E(\hat{\theta}) = \sum_{n=0}^{N-1} a_n E(x[n]) = \theta.$$

□ برای اینکه این شرط رعایت شود، لزوماً باید داشته باشیم:

$$E(x[n]) = s[n]\theta$$

□ چون اگر $E(x[n])$ بر حسب θ خطی نباشد، شرط بدون بایاس بودن محقق نخواهد شد.

□ بنابراین باید تابع زیر را با شرط پیش گفته کمینه کنیم:

$$\text{var}(\hat{\theta}) = \mathbf{a}^T \mathbf{C} \mathbf{a}$$

□ شرط بدون بایاس بودن منجر می شود به:

$$\sum_{n=0}^{N-1} a_n E(x[n]) = \theta$$

$$\sum_{n=0}^{N-1} a_n s[n] \theta = \theta$$

$$\sum_{n=0}^{N-1} a_n s[n] = 1$$

$$\mathbf{a}^T \mathbf{s} = 1$$

□ مساله ریاضی که باید حل شود:

$$\min_{\mathbf{a}} \mathbf{a}^T \mathbf{C} \mathbf{a} \quad \text{subject to: } \mathbf{a}^T \mathbf{s} = 1$$

□ حل مساله از روش لاگرانژ:

$$\mathbf{a}_{\text{opt}} = \frac{\mathbf{C}^{-1}\mathbf{s}}{\mathbf{s}^T\mathbf{C}^{-1}\mathbf{s}}$$

□ بنابراین BLUE به دست می آید:

$$\hat{\theta} = \frac{\mathbf{s}^T\mathbf{C}^{-1}\mathbf{x}}{\mathbf{s}^T\mathbf{C}^{-1}\mathbf{s}}$$

□ واریانس کمینه:

$$\text{var}(\hat{\theta}) = \frac{1}{\mathbf{s}^T\mathbf{C}^{-1}\mathbf{s}}.$$

□ شرط بدون بایاس بودن برقرار است:

$$\begin{aligned} E(\hat{\theta}) &= \frac{\mathbf{s}^T\mathbf{C}^{-1}E(\mathbf{x})}{\mathbf{s}^T\mathbf{C}^{-1}\mathbf{s}} \\ &= \frac{\mathbf{s}^T\mathbf{C}^{-1}\theta\mathbf{s}}{\mathbf{s}^T\mathbf{C}^{-1}\mathbf{s}} \\ &= \theta. \end{aligned}$$

□ آنچه لازم است برای تحقق BLUE بدانیم:

□ s یا همان میانگین اسکیل شده داده،

□ C یا ماتریس کوواریانس داده

□ به PDF نیاز نداریم.

□ مثال ۱: سطح DC در نویز سفید

$$x[n] = A + w[n] \quad n = 0, 1, \dots, N - 1$$

where $w[n]$ is white noise with variance σ^2 (and of unspecified PDF)

□ دقت کنید که در این مثال PDF نویز را نداریم (دلخواه است)، فقط میانگین و واریانس را داریم.

$$E(x[n]) = A$$

□ داریم:

$$s[n] = 1 \text{ and therefore } s = 1$$

□ بنابراین:

$$\begin{aligned}\hat{A} &= \frac{\mathbf{1}^T \frac{1}{\sigma^2} \mathbf{I} \mathbf{x}}{\mathbf{1}^T \frac{1}{\sigma^2} \mathbf{I} \mathbf{1}} \\ &= \frac{1}{N} \sum_{n=0}^{N-1} x[n] = \bar{x}\end{aligned}$$

$$\begin{aligned}\text{var}(\hat{A}) &= \frac{1}{\mathbf{1}^T \frac{1}{\sigma^2} \mathbf{1}} \\ &= \frac{\sigma^2}{N}.\end{aligned}$$

□ و در نهایت:

□ دقت کنید که اگر نویز گوسی باشد، BLUE با MVUE برابر است، ولی در غیر این صورت BLUE یک راه حل زیربهبینه است.

□ مثال ۲: سطح DC در نویز ناهمبسته: $\text{var}(w[n]) = \sigma_n^2$

□ مانند قبل داریم: $\mathbf{s} = \mathbf{1}$

□ بنابراین:

$$\hat{A} = \frac{\mathbf{1}^T \mathbf{C}^{-1} \mathbf{x}}{\mathbf{1}^T \mathbf{C}^{-1} \mathbf{1}}$$

$$\text{var}(\hat{A}) = \frac{1}{\mathbf{1}^T \mathbf{C}^{-1} \mathbf{1}}.$$

□ توضیح: با توجه به اینکه:

$$\mathbf{C} = \begin{bmatrix} \sigma_0^2 & 0 & \dots & 0 \\ 0 & \sigma_1^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_{N-1}^2 \end{bmatrix},$$

می توان به دست آورد:

$$\hat{A} = \frac{\sum_{n=0}^{N-1} \frac{x[n]}{\sigma_n^2}}{\sum_{n=0}^{N-1} \frac{1}{\sigma_n^2}}$$

یعنی در اینجا BLUE همان میانگین گیری وزن دار داده هاست. هر داده ای که نویز بیشتری داشته باشد با وزن کمتر در میانگین شرکت می کند.

ضمنا داریم:

$$\text{var}(\hat{A}) = \frac{1}{\sum_{n=0}^{N-1} \frac{1}{\sigma_n^2}}.$$

تعمیم به حالت برداری

15

□ در اینجا p پارامتر داریم که می خواهیم به صورت خطی تخمین بزنیم:

$$\hat{\theta}_i = \sum_{n=0}^{N-1} a_{in} x[n] \quad i = 1, 2, \dots, p$$

$$\hat{\boldsymbol{\theta}} = \mathbf{A}\mathbf{x}$$

□ شرط بدون بایاس بودن:

$$E(\hat{\theta}_i) = \sum_{n=0}^{N-1} a_{in} E(x[n]) = \theta_i \quad i = 1, 2, \dots, p$$

$$E(\hat{\boldsymbol{\theta}}) = \mathbf{A}E(\mathbf{x}) = \boldsymbol{\theta}.$$

□ برای اینکه شرط بدون بایاس بودن محقق شود باید داشته باشیم:

$$E(\mathbf{x}) = \mathbf{H}\boldsymbol{\theta} \quad \longrightarrow \quad \mathbf{A}\mathbf{H} = \mathbf{I}.$$

تعمیم به حالت برداری

16

□ با فرض:

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \vdots \\ \mathbf{a}_p^T \end{bmatrix} \quad \mathbf{H} = [\mathbf{h}_1 \quad \mathbf{h}_2 \quad \dots \quad \mathbf{h}_p].$$

□ شرط پیش گفته منجر می شود به:

$$\mathbf{a}_i^T \mathbf{h}_j = \delta_{ij} \quad i = 1, 2, \dots, p; j = 1, 2, \dots, p.$$

□ بنابراین باید مقدار واریانس زیر با شرط بالا کمینه شود:

$$\text{var}(\hat{\theta}_i) = \mathbf{a}_i^T \mathbf{C} \mathbf{a}_i$$

□ نتیجه:

$$\hat{\boldsymbol{\theta}} = (\mathbf{H}^T \mathbf{C}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{C}^{-1} \mathbf{x}$$

$$\mathbf{C}_{\hat{\boldsymbol{\theta}}} = (\mathbf{H}^T \mathbf{C}^{-1} \mathbf{H})^{-1}.$$

Theorem 6.1 (Gauss-Markov Theorem) *If the data are of the general linear model form*

$$\mathbf{x} = \mathbf{H}\boldsymbol{\theta} + \mathbf{w} \quad (6.18)$$

where $\mathbf{H}$ is a known $N \times p$ matrix, $\boldsymbol{\theta}$ is a $p \times 1$ vector of parameters to be estimated, and $\mathbf{w}$ is a $N \times 1$ noise vector with zero mean and covariance $\mathbf{C}$ (the PDF of $\mathbf{w}$ is otherwise arbitrary), then the BLUE of $\boldsymbol{\theta}$ is

$$\hat{\boldsymbol{\theta}} = (\mathbf{H}^T \mathbf{C}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{C}^{-1} \mathbf{x} \quad (6.19)$$

and the minimum variance of $\hat{\theta}_i$ is

$$\text{var}(\hat{\theta}_i) = [(\mathbf{H}^T \mathbf{C}^{-1} \mathbf{H})^{-1}]_{ii}. \quad (6.20)$$

In addition, the covariance matrix of $\hat{\boldsymbol{\theta}}$ is

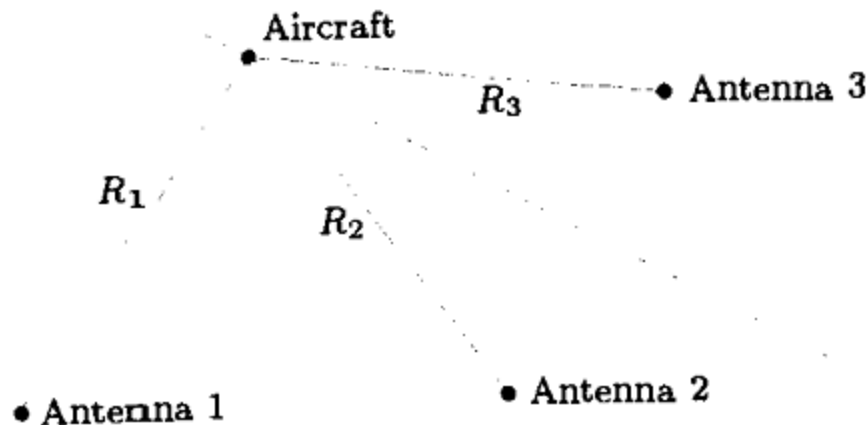
$$\mathbf{C}_{\hat{\boldsymbol{\theta}}} = (\mathbf{H}^T \mathbf{C}^{-1} \mathbf{H})^{-1}. \quad (6.21)$$

تعمیم به حالت برداری

18

□ مثال: مکان یابی منبع (source localization)

□ در بسیاری از کاربردها علاقه مندیم با اندازه گیری زمان ورود سیگنال از یک منبع، مکان آن منبع را تخمین بزنیم.



$$\left. \begin{array}{l} t_2 - t_1 = 0 \\ t_3 - t_2 = 0 \end{array} \right\} \Rightarrow R_1 = R_2 = R_3$$

t_i = Time of received signal (emitted by aircraft) at antenna i

تعمیم به حالت برداری

19

□ در یک هندسه با N آنتن، زمان دریافت سیگنال به وسیله هر آنتن به

صورت زیر خواهد بود:

$$t_i = T_0 + R_i/c + \epsilon_i \quad i = 0, 1, \dots, N - 1$$

$$R_i = \sqrt{(x_s - x_i)^2 + (y_s - y_i)^2}.$$

□ مشخص است که جایگذاری R_i در معادله بالا باعث غیر خطی شدن روابط خواهد شد. به همین علت می توان حول یک نقطه نامی (مثلا محل هدف در اندازه گیری قبل) روابط را به صورت خطی تقریب زد:

$$R_i \approx R_{n_i} + \frac{x_n - x_i}{R_{n_i}} \delta x_s + \frac{y_n - y_i}{R_{n_i}} \delta y_s.$$

□ آنچه اکنون به دنبال تخمین آن هستیم:

$$\theta = [(x_s - x_n) \ (y_s - y_n)]^T = [\delta x_s \ \delta y_s]^T$$

تعمیم به حالت برداری

20

□ بنابراین:

$$t_i = T_0 + \frac{R_{n_i}}{c} + \frac{x_n - x_i}{R_{n_i}c} \delta x_s + \frac{y_n - y_i}{R_{n_i}c} \delta y_s + \epsilon_i$$

□ با توجه به اینکه:

$$\frac{x_n - x_i}{R_{n_i}} = \cos \alpha_i \quad \frac{y_n - y_i}{R_{n_i}} = \sin \alpha_i,$$

□ می توان نوشت:

$$t_i = T_0 + \frac{R_{n_i}}{c} + \frac{\cos \alpha_i}{c} \delta x_s + \frac{\sin \alpha_i}{c} \delta y_s + \epsilon_i.$$

□ با تعریف: $\tau_i = t_i - \frac{R_{n_i}}{c}$

□ داریم:

$$\tau_i = T_0 + \frac{\cos \alpha_i}{c} \delta x_s + \frac{\sin \alpha_i}{c} \delta y_s + \epsilon_i$$

تعمیم به حالت برداری

21

□ برای حذف پارامتر نامعلوم τ_0 می توان از تفاضل داده ها استفاده کرد
(تعداد داده ها یکی کم می شود):

$$\xi_1 = \tau_1 - \tau_0$$

$$\xi_2 = \tau_2 - \tau_1$$

$$\vdots$$

$$\xi_{N-1} = \tau_{N-1} - \tau_{N-2}.$$

□ که:

$$\xi_i = \frac{1}{c}(\cos \alpha_i - \cos \alpha_{i-1})\delta x_s + \frac{1}{c}(\sin \alpha_i - \sin \alpha_{i-1})\delta y_s + \epsilon_i - \epsilon_{i-1}$$

$$i = 1, 2, \dots, N - 1.$$

□ در نهایت با توجه به روابط فوق خواهیم داشت:

تعمیم به حالت برداری

22

$$\boldsymbol{\theta} = \begin{bmatrix} \delta x_s & \delta y_s \end{bmatrix}^T$$

$$\mathbf{H} = \frac{1}{c} \begin{bmatrix} \cos \alpha_1 - \cos \alpha_0 & \sin \alpha_1 - \sin \alpha_0 \\ \cos \alpha_2 - \cos \alpha_1 & \sin \alpha_2 - \sin \alpha_1 \\ \vdots & \vdots \\ \cos \alpha_{N-1} - \cos \alpha_{N-2} & \sin \alpha_{N-1} - \sin \alpha_{N-2} \end{bmatrix}$$

$$\mathbf{w} = \begin{bmatrix} \epsilon_1 - \epsilon_0 \\ \epsilon_2 - \epsilon_1 \\ \vdots \\ \epsilon_{N-1} - \epsilon_{N-2} \end{bmatrix}.$$

□ ضمناً داریم:

$$\mathbf{w} = \underbrace{\begin{bmatrix} -1 & 1 & 0 & 0 & \dots & 0 \\ 0 & -1 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & -1 & 1 \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} \epsilon_0 \\ \epsilon_1 \\ \vdots \\ \epsilon_{N-1} \end{bmatrix}}_{\boldsymbol{\epsilon}}, \quad \Rightarrow \quad \mathbf{C} = E[\mathbf{A}\boldsymbol{\epsilon}\boldsymbol{\epsilon}^T\mathbf{A}^T] = \sigma^2\mathbf{A}\mathbf{A}^T$$

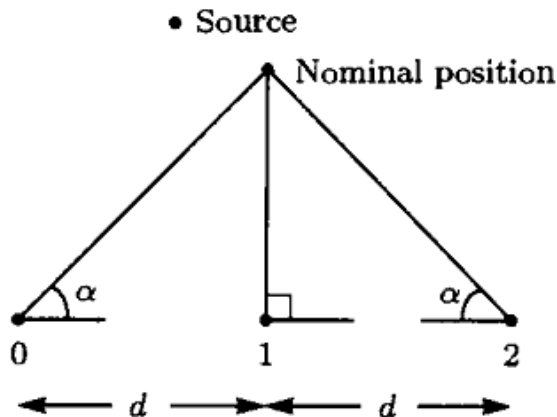
تعمیم به حالت برداری

□ در نهایت بهترین تخمینگر خطی (BLUE) از «تغییر مکان» هدف:

$$\begin{aligned}\hat{\theta} &= (\mathbf{H}^T \mathbf{C}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{C}^{-1} \xi \\ &= [\mathbf{H}^T (\mathbf{A} \mathbf{A}^T)^{-1} \mathbf{H}]^{-1} \mathbf{H}^T (\mathbf{A} \mathbf{A}^T)^{-1} \xi\end{aligned}$$

$$\mathbf{C}_{\hat{\theta}} = \sigma^2 [\mathbf{H}^T (\mathbf{A} \mathbf{A}^T)^{-1} \mathbf{H}]^{-1}.$$

□ به عنوان مثال در حالت سه آنتنی زیر داریم:



$$\mathbf{H} = \frac{1}{c} \begin{bmatrix} -\cos \alpha & 1 - \sin \alpha \\ -\cos \alpha & -(1 - \sin \alpha) \end{bmatrix}$$

$$\mathbf{A} = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}.$$

تعمیم به حالت برداری

□ ماتریس کوواریانس تخمین:

$$\mathbf{C}_{\hat{\theta}} = \sigma^2 c^2 \begin{bmatrix} \frac{1}{2 \cos^2 \alpha} & 0 \\ 0 & \frac{3/2}{(1 - \sin \alpha)^2} \end{bmatrix}.$$

□ یعنی برای تخمین بهتر باید مقدار زاویه α کوچک باشد که به معنی برد کمتر یا گستره آرایه آنتنی بزرگتر است.