



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

تئوری تخمین

ESTIMATION THEORY

نیم سال اول سال تحصیلی ۰۴-۰۵

- معرفی
- بیان قضیه
- حالت کلی برای سیگنال های آلوده به نویز گوسی
- تبدیل پارامترها
- تعمیم به حالت برداری

- حد کرامر-راو (CRLB) یک حد پایین برای واریانس همه تخمینگرهای بدون بایاس ارائه می کند.
- اگر واریانس یک تخمینگر بدون بایاس مساوی با این حد باشد، به این معنی است که در بین همه تخمینگرهای بدون بایاس، این تخمینگر بهترین است.
- فاصله هر تخمینگر بدون بایاس با این حد به عنوان معیاری از کیفیت آن تخمینگر در نظر گرفته می شود.

Theorem 3.1 (Cramer-Rao Lower Bound - Scalar Parameter) *It is assumed that the PDF $p(\mathbf{x}; \theta)$ satisfies the “regularity” condition*

$$E \left[\frac{\partial \ln p(\mathbf{x}; \theta)}{\partial \theta} \right] = 0 \quad \text{for all } \theta$$

where the expectation is taken with respect to $p(\mathbf{x}; \theta)$. Then, the variance of any unbiased estimator $\hat{\theta}$ must satisfy

$$\text{var}(\hat{\theta}) \geq \frac{1}{-E \left[\frac{\partial^2 \ln p(\mathbf{x}; \theta)}{\partial \theta^2} \right]} \quad (3.6)$$

where the derivative is evaluated at the true value of θ and the expectation is taken with respect to $p(\mathbf{x}; \theta)$. Furthermore, an unbiased estimator may be found that attains the bound for all θ if and only if

$$\frac{\partial \ln p(\mathbf{x}; \theta)}{\partial \theta} = I(\theta)(g(\mathbf{x}) - \theta) \quad (3.7)$$

for some functions g and I . That estimator, which is the MVU estimator, is $\hat{\theta} = g(\mathbf{x})$, and the minimum variance is $1/I(\theta)$.

□ مثال ۱

$$x[0] = A + w[0] \quad \text{where } w[0] \sim \mathcal{N}(0, \sigma^2)$$

$$p(x[0]; A) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[-\frac{1}{2\sigma^2} (x[0] - A)^2 \right]$$

□ شرط regularity برقرار است:

$$\ln p(x[0]; A) = -\ln \sqrt{2\pi\sigma^2} - \frac{1}{2\sigma^2} (x[0] - A)^2$$

$$\frac{\partial \ln p(x[0]; A)}{\partial A} = \frac{1}{\sigma^2} (x[0] - A)$$

$$E \left\{ \frac{\partial \ln(p(x[0]; A))}{\partial A} \right\} = \frac{1}{\sigma^2} (E\{x[0]\} - A) = 0$$

□ ادامه مثال ۱

$$-\frac{\partial^2 \ln p(x[0]; A)}{\partial A^2} = \frac{1}{\sigma^2} \quad \Rightarrow \quad \text{var}(\hat{A}) \geq \frac{1}{-\frac{\partial^2 \ln(p(x[0]; A))}{\partial A^2}} = \sigma^2$$

□ ضمناً طبق شرط لازم و کافی در صورتی که داشته باشیم:

$$\frac{\partial \ln p(\mathbf{x}; \theta)}{\partial \theta} = I(\theta)(g(\mathbf{x}) - \theta)$$

□ آنگاه مقدار تخمینگر MVU همان $g(\mathbf{x})$ خواهد بود. با مقایسه با رابطه قبل داریم:

$$\begin{aligned} \theta &= A \\ I(\theta) &= \frac{1}{\sigma^2} \\ g(x[0]) &= x[0] \end{aligned}$$

□ ادامه مثال ۱:

□ بنابراین:

$\hat{A} = g(x[0]) = x[0]$ is the MVU estimator.

$$\text{var}(\hat{A}) = \sigma^2 = 1/I(\theta)$$

□ مثال ۲:

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

$w[n]$ is WGN with variance σ^2

□ ادامه مثال ۲

$$p(\mathbf{x}; A) = \prod_{n=0}^{N-1} \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[-\frac{1}{2\sigma^2} (x[n] - A)^2 \right]$$

$$= \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} \exp \left[-\frac{1}{2\sigma^2} \sum_{n=0}^{N-1} (x[n] - A)^2 \right].$$

$$\frac{\partial \ln p(\mathbf{x}; A)}{\partial A} = \frac{\partial}{\partial A} \left[-\ln[(2\pi\sigma^2)^{\frac{N}{2}}] - \frac{1}{2\sigma^2} \sum_{n=0}^{N-1} (x[n] - A)^2 \right]$$

$$= \frac{1}{\sigma^2} \sum_{n=0}^{N-1} (x[n] - A)$$

$$= \frac{N}{\sigma^2} (\bar{x} - A)$$

$$\frac{\partial^2 \ln p(\mathbf{x}; A)}{\partial A^2} = -\frac{N}{\sigma^2}$$



$$\text{var}(\hat{A}) \geq \frac{\sigma^2}{N}$$

□ ادامه مثال ۲

$$\frac{\partial \ln p(\mathbf{x}; \theta)}{\partial \theta} = I(\theta)(g(\mathbf{x}) - \theta)$$

□ ضمناً با توجه به رابطه روبرو:

$$\frac{\partial \ln p(\mathbf{x}; A)}{\partial A} = \frac{N}{\sigma^2}(\bar{x} - A)$$

و مقایسه آن با رابطه:

مشخص است که تخمینگر MVU همان مقدار میانگین است و کمترین واریانس ممکن برابر با σ^2/N است.

□ مثال ۳. فرض کنید می خواهیم فاز یک سیگنال سینوسی مخلوط با نویز گوسی را تخمین بزنیم. دامنه و فرکانس معلوم فرض می شوند.

$$x[n] = A \cos(2\pi f_0 n + \phi) + w[n] \quad n = 0, 1, \dots, N-1.$$

□ داریم:

$$p(\mathbf{x}; \phi) = \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{n=0}^{N-1} [x[n] - A \cos(2\pi f_0 n + \phi)]^2 \right\}.$$

□ مشتق گیری از تابع log-likelihood:

$$\begin{aligned} \frac{\partial \ln p(\mathbf{x}; \phi)}{\partial \phi} &= -\frac{1}{\sigma^2} \sum_{n=0}^{N-1} [x[n] - A \cos(2\pi f_0 n + \phi)] A \sin(2\pi f_0 n + \phi) \\ &= -\frac{A}{\sigma^2} \sum_{n=0}^{N-1} [x[n] \sin(2\pi f_0 n + \phi) - \frac{A}{2} \sin(4\pi f_0 n + 2\phi)] \end{aligned}$$

□ ادامه مثال ۳

$$\frac{\partial^2 \ln p(\mathbf{x}; \phi)}{\partial \phi^2} = -\frac{A}{\sigma^2} \sum_{n=0}^{N-1} [x[n] \cos(2\pi f_0 n + \phi) - A \cos(4\pi f_0 n + 2\phi)].$$

□ امید ریاضی

$$\begin{aligned} -E \left[\frac{\partial^2 \ln p(\mathbf{x}; \phi)}{\partial \phi^2} \right] &= \frac{A}{\sigma^2} \sum_{n=0}^{N-1} [A \cos^2(2\pi f_0 n + \phi) - A \cos(4\pi f_0 n + 2\phi)] \\ &= \frac{A^2}{\sigma^2} \sum_{n=0}^{N-1} \left[\frac{1}{2} + \frac{1}{2} \cos(4\pi f_0 n + 2\phi) - \cos(4\pi f_0 n + 2\phi) \right] \\ &\approx \frac{NA^2}{2\sigma^2} \end{aligned}$$

□ زیرا برای f_0 مخالف صفر و $1/2$:

$$\frac{1}{N} \sum_{n=0}^{N-1} \cos(4\pi f_0 n + 2\phi) \approx 0$$

□ ادامه مثال ۳

$$\text{var}(\hat{\phi}) \geq \frac{2\sigma^2}{NA^2}.$$

□ بنابراین:

□ ضمناً با توجه به مشتق تابع $\log\text{-likelihood}$ مشخص است که تخمینگر فازی که بدون بایاس بوده و واریانس آن به حد کرامر-راو برسد وجود ندارد.

□ دقت کنید که در این حالت هم ممکن است یک تخمینگر MVU وجود داشته باشد.

□ در حال حاضر ما نمی دانیم که آیا یک تخمینگر MVU وجود دارد؟ و اینکه چگونه می توان آن را به دست آورد؟

حالت کلی سیگنال آغشته به نویز گوسی

13

□ در اینجا به طور کلی داریم:

$$x[n] = s[n; \theta] + w[n] \quad n = 0, 1, \dots, N-1.$$

□ تابع likelihood:

$$p(\mathbf{x}; \theta) = \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{n=0}^{N-1} (x[n] - s[n; \theta])^2 \right\}.$$

□ مشتق گیری:

$$\frac{\partial \ln p(\mathbf{x}; \theta)}{\partial \theta} = \frac{1}{\sigma^2} \sum_{n=0}^{N-1} (x[n] - s[n; \theta]) \frac{\partial s[n; \theta]}{\partial \theta}$$

$$\frac{\partial^2 \ln p(\mathbf{x}; \theta)}{\partial \theta^2} = \frac{1}{\sigma^2} \sum_{n=0}^{N-1} \left\{ (x[n] - s[n; \theta]) \frac{\partial^2 s[n; \theta]}{\partial \theta^2} - \left(\frac{\partial s[n; \theta]}{\partial \theta} \right)^2 \right\}.$$

حالت کلی سیگنال آغشته به نویز گوسی

14

$$E \left(\frac{\partial^2 \ln p(\mathbf{x}; \theta)}{\partial \theta^2} \right) = -\frac{1}{\sigma^2} \sum_{n=0}^{N-1} \left(\frac{\partial s[n; \theta]}{\partial \theta} \right)^2 \quad \square \text{ گرفتن امید:}$$

$$\text{var}(\hat{\theta}) \geq \frac{\sigma^2}{\sum_{n=0}^{N-1} \left(\frac{\partial s[n; \theta]}{\partial \theta} \right)^2}.$$

$\square$ و در نهایت:

$\square$ مثال ۴: تخمین فرکانس سیگنال سینوسی

$$s[n; f_0] = A \cos(2\pi f_0 n + \phi) \quad 0 < f_0 < \frac{1}{2}$$

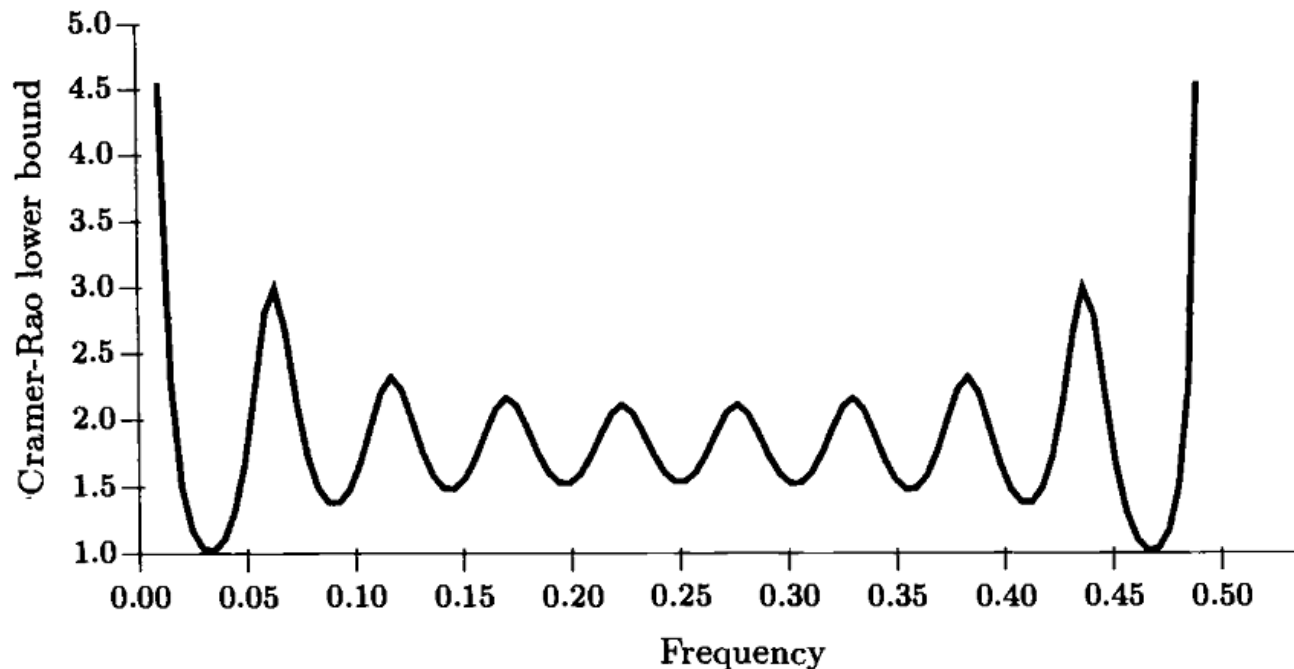
حالت کلی سیگنال آغشته به نویز گوسی

15

□ ادامه مثال ۴

□ طبق فرمول صفحه قبل:

$$\text{var}(\hat{f}_0) \geq \frac{\sigma^2}{A^2 \sum_{n=0}^{N-1} [2\pi n \sin(2\pi f_0 n + \phi)]^2}.$$



$$\frac{A^2}{\sigma^2} = 1$$

$$N = 10$$

$$\phi = 0$$

تبدیل پارامترها

16

□ در برخی موارد علاقه مندیم که تابعی از پارامتر مورد نظر را تخمین بزنیم. به طور کلی اگر بخواهیم مقدار α را (که تابعی از پارامتر مورد نظر به صورت $\alpha = g(\theta)$ است) تخمین بزنیم:

$$\text{var}(\hat{\alpha}) \geq \frac{\left(\frac{\partial g}{\partial \theta}\right)^2}{-E\left[\frac{\partial^2 \ln p(\mathbf{x}; \theta)}{\partial \theta^2}\right]}.$$

□ مثال ۵: فرض کنید در مثال اندازه گیری ولتاژ، می خواهیم مقدار مجذور ولتاژ را تخمین بزنیم.

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

$$\alpha = g(A) = A^2$$

□ ادامه مثال ۵

$$\text{var}(\widehat{A^2}) \geq \frac{(2A)^2}{N/\sigma^2} = \frac{4A^2\sigma^2}{N}.$$

□ ممکن است به نظر برسد چون $\bar{x}$ تخمین کارامدی از A است، پس $\bar{x}^2$ نیز تخمین کارامدی از A^2 است.

□ این موضوع را با بررسی بایاس بررسی می کنیم:

$$\begin{aligned} E(\bar{x}^2) &= E^2(\bar{x}) + \text{var}(\bar{x}) = A^2 + \frac{\sigma^2}{N} \\ &\neq A^2. \end{aligned}$$

□ می بینیم که تخمینگر $\bar{x}^2$ برای تخمین A^2 بایاس دارد.

تبدیل پارامترها

18

□ بنابراین به طور کلی کارایی یک تخمینگر با اعمال تبدیل غیرخطی از دست می رود.

□ در مورد تبدیل خطی، کارایی تخمینگر حفظ می شود: $g(\theta) = a\theta + b$.
 □ انتخاب تخمینگر:

$$\widehat{g(\theta)} = g(\hat{\theta}) = a\hat{\theta} + b$$

□ بررسی بایاس:

$$E(a\hat{\theta} + b) = aE(\hat{\theta}) + b = a\theta + b = g(\theta)$$

□ بررسی کارایی:

$$\text{var}(\widehat{g(\theta)}) \geq \frac{\left(\frac{\partial g}{\partial \theta}\right)^2}{I(\theta)} = \left(\frac{\partial g(\theta)}{\partial \theta}\right)^2 \text{var}(\hat{\theta}) = a^2 \text{var}(\hat{\theta}).$$

□ از طرفی:

$$\text{var}(\widehat{g(\theta)}) = \text{var}(a\hat{\theta} + b) = a^2 \text{var}(\hat{\theta})$$

تبدیل پارامترها

□ در مورد تبدیل های غیر خطی در صورتی که **داده گسترده** در اختیار داشته باشیم، کارایی به صورت تقریبی (مجانبی) حفظ می شود.

□ به عنوان مثال در مورد تخمین A^2 با استفاده از تخمینگر $\bar{x}^2$ داشتیم:

$$E(\bar{x}^2) = E^2(\bar{x}) + \text{var}(\bar{x}) = A^2 + \frac{\sigma^2}{N} \neq A^2.$$

$$\text{var}(\bar{x}^2) = \frac{4A^2\sigma^2}{N} + \frac{2\sigma^4}{N^2}. \quad \text{as } N \rightarrow \infty, \text{ the variance approaches } 4A^2\sigma^2/N$$

□ می بینیم که اگر مقدار N خیلی بزرگ باشد، بدون بایاس بودن و کارایی تخمینگر به صورت مجانبی حفظ می شود.

تعمیم به حالت برداری

20

□ فرض کنید می خواهیم بردار پارامتر زیر را تخمین بزنیم:

$$\boldsymbol{\theta} = [\theta_1 \theta_2 \dots \theta_p]^T$$

□ در این صورت حد کرامر-راو برای تخمین پارامتر $\boldsymbol{\theta}$ به صورت زیر خواهد بود:

$$\text{var}(\hat{\theta}_i) \geq [\mathbf{I}^{-1}(\boldsymbol{\theta})]_{ii}$$

□ که ماتریس $\mathbf{I}(\boldsymbol{\theta})$ ماتریس اطلاعات فشر نامیده می شود و به صورت زیر به دست می آید:

$$[\mathbf{I}(\boldsymbol{\theta})]_{ij} = -E \left[\frac{\partial^2 \ln p(\mathbf{x}; \boldsymbol{\theta})}{\partial \theta_i \partial \theta_j} \right]$$

تعمیم به حالت برداری

□ مثال ۶. فرض کنید در مثال ولتاژ DC، علاوه بر مقدار ولتاژ می خواهیم مقدار واریانس نویز را هم تخمین بزنیم.

$$\boldsymbol{\theta} = [A \sigma^2]^T$$

$$\mathbf{I}(\boldsymbol{\theta}) = \begin{bmatrix} -E \left[\frac{\partial^2 \ln p(\mathbf{x}; \boldsymbol{\theta})}{\partial A^2} \right] & -E \left[\frac{\partial^2 \ln p(\mathbf{x}; \boldsymbol{\theta})}{\partial A \partial \sigma^2} \right] \\ -E \left[\frac{\partial^2 \ln p(\mathbf{x}; \boldsymbol{\theta})}{\partial \sigma^2 \partial A} \right] & -E \left[\frac{\partial^2 \ln p(\mathbf{x}; \boldsymbol{\theta})}{\partial \sigma^2{}^2} \right] \end{bmatrix}.$$

$$\ln p(\mathbf{x}; \boldsymbol{\theta}) = -\frac{N}{2} \ln 2\pi - \frac{N}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} \sum_{n=0}^{N-1} (x[n] - A)^2.$$

تعمیم به حالت برداری

22

□ ادامه مثال ۶

$$\frac{\partial \ln p(\mathbf{x}; \boldsymbol{\theta})}{\partial A} = \frac{1}{\sigma^2} \sum_{n=0}^{N-1} (x[n] - A)$$

$$\frac{\partial \ln p(\mathbf{x}; \boldsymbol{\theta})}{\partial \sigma^2} = -\frac{N}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{n=0}^{N-1} (x[n] - A)^2$$

$$\frac{\partial^2 \ln p(\mathbf{x}; \boldsymbol{\theta})}{\partial A^2} = -\frac{N}{\sigma^2}$$

$$\frac{\partial^2 \ln p(\mathbf{x}; \boldsymbol{\theta})}{\partial A \partial \sigma^2} = -\frac{1}{\sigma^4} \sum_{n=0}^{N-1} (x[n] - A)$$

$$\frac{\partial^2 \ln p(\mathbf{x}; \boldsymbol{\theta})}{\partial \sigma^2{}^2} = \frac{N}{2\sigma^4} - \frac{1}{\sigma^6} \sum_{n=0}^{N-1} (x[n] - A)^2.$$

تعمیم به حالت برداری

23

□ ادامه مثال ۶

□ با منفی کردن و گرفتن امید:

$$\mathbf{I}(\boldsymbol{\theta}) = \begin{bmatrix} \frac{N}{\sigma^2} & 0 \\ 0 & \frac{N}{2\sigma^4} \end{bmatrix}.$$

□ در نتیجه:

$$\begin{aligned} \text{var}(\hat{A}) &\geq \frac{\sigma^2}{N} \\ \text{var}(\hat{\sigma}^2) &\geq \frac{2\sigma^4}{N}. \end{aligned}$$

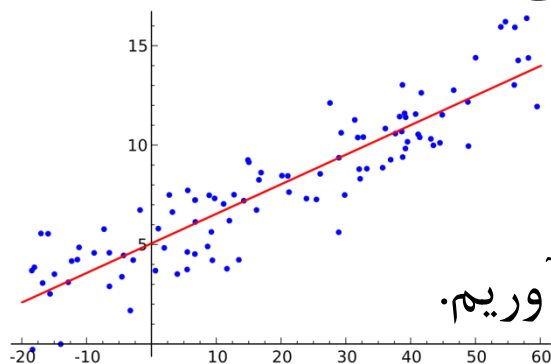
□ همانطور که می بینیم حد واریانس تخمین A در حالتی که σ^2 معلوم است نیز همین است.

□ این مساله به خاطر قطری بودن ماتریس است و همیشه برقرار نیست.

تعمیم به حالت برداری

24

□ مثال ۷. برازش خط: line fitting (رگرسیون خطی)



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

where $w[n]$ is WGN

□ می خواهیم CRLB برای ضرایب A و B را به دست آوریم.

□ داریم:

$$\boldsymbol{\theta} = [A \ B]^T$$

□ بنابراین:

$$\mathbf{I}(\boldsymbol{\theta}) = \begin{bmatrix} -E \left[\frac{\partial^2 \ln p(\mathbf{x}; \boldsymbol{\theta})}{\partial A^2} \right] & -E \left[\frac{\partial^2 \ln p(\mathbf{x}; \boldsymbol{\theta})}{\partial A \partial B} \right] \\ -E \left[\frac{\partial^2 \ln p(\mathbf{x}; \boldsymbol{\theta})}{\partial B \partial A} \right] & -E \left[\frac{\partial^2 \ln p(\mathbf{x}; \boldsymbol{\theta})}{\partial B^2} \right] \end{bmatrix}.$$

تعمیم به حالت برداری

□ ادامه مثال ۷. تابع likelihood و مشتقات:

$$p(\mathbf{x}; \boldsymbol{\theta}) = \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{n=0}^{N-1} (x[n] - A - Bn)^2 \right\}$$

$$\frac{\partial \ln p(\mathbf{x}; \boldsymbol{\theta})}{\partial A} = \frac{1}{\sigma^2} \sum_{n=0}^{N-1} (x[n] - A - Bn)$$

$$\frac{\partial \ln p(\mathbf{x}; \boldsymbol{\theta})}{\partial B} = \frac{1}{\sigma^2} \sum_{n=0}^{N-1} (x[n] - A - Bn)n$$

$$\frac{\partial^2 \ln p(\mathbf{x}; \boldsymbol{\theta})}{\partial A^2} = -\frac{N}{\sigma^2}$$

$$\frac{\partial^2 \ln p(\mathbf{x}; \boldsymbol{\theta})}{\partial A \partial B} = -\frac{1}{\sigma^2} \sum_{n=0}^{N-1} n$$

$$\frac{\partial^2 \ln p(\mathbf{x}; \boldsymbol{\theta})}{\partial B^2} = -\frac{1}{\sigma^2} \sum_{n=0}^{N-1} n^2$$

تعمیم به حالت برداری

26

□ ادامه مثال ۷. ماتریس اطلاعات فشر:

$$\mathbf{I}(\boldsymbol{\theta}) = \frac{1}{\sigma^2} \begin{bmatrix} N & \sum_{n=0}^{N-1} n \\ \sum_{n=0}^{N-1} n & \sum_{n=0}^{N-1} n^2 \end{bmatrix} = \frac{1}{\sigma^2} \begin{bmatrix} N & \frac{N(N-1)}{2} \\ \frac{N(N-1)}{2} & \frac{N(N-1)(2N-1)}{6} \end{bmatrix}$$

□ محاسبه معکوس:

$$\mathbf{I}^{-1}(\boldsymbol{\theta}) = \sigma^2 \begin{bmatrix} \frac{2(2N-1)}{N(N+1)} & -\frac{6}{N(N+1)} \\ -\frac{6}{N(N+1)} & \frac{12}{N(N^2-1)} \end{bmatrix}.$$

$$\begin{aligned} \text{var}(\hat{A}) &\geq \frac{2(2N-1)\sigma^2}{N(N+1)} \\ \text{var}(\hat{B}) &\geq \frac{12\sigma^2}{N(N^2-1)}. \end{aligned}$$

□ بنابراین:

تعمیم به حالت برداری

□ ادامه مثال ۷

□ اگر مقدار B را معلوم فرض کنیم داریم:

$$\text{var}(\hat{A}) \geq -\frac{1}{E \left[\frac{\partial^2 \ln p(\mathbf{x}; A)}{\partial A^2} \right]} = \frac{\sigma^2}{N}$$

□ همانطور که می بینیم حد CRLB برای تخمین A در حالتی که B معلوم است، کمتر از حالتی است که B مجهول است.

□ نکته دیگر این است که با گسترش داده، واریانس تخمین B و A به ترتیب با $\frac{1}{N}$ و $\frac{1}{N^3}$ متناسب است. یعنی تخمین B «آسان تر» است.

Theorem 3.2 (Cramer-Rao Lower Bound - Vector Parameter) *It is assumed that the PDF $p(\mathbf{x}; \boldsymbol{\theta})$ satisfies the “regularity” conditions*

$$E \left[\frac{\partial \ln p(\mathbf{x}; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right] = \mathbf{0} \quad \text{for all } \boldsymbol{\theta}$$

where the expectation is taken with respect to $p(\mathbf{x}; \boldsymbol{\theta})$. Then, the covariance matrix of any unbiased estimator $\hat{\boldsymbol{\theta}}$ satisfies

$$\mathbf{C}_{\hat{\boldsymbol{\theta}}} - \mathbf{I}^{-1}(\boldsymbol{\theta}) \geq \mathbf{0} \quad (3.24)$$

where $\geq \mathbf{0}$ is interpreted as meaning that the matrix is positive semidefinite. The Fisher information matrix $\mathbf{I}(\boldsymbol{\theta})$ is given as

$$[\mathbf{I}(\boldsymbol{\theta})]_{ij} = -E \left[\frac{\partial^2 \ln p(\mathbf{x}; \boldsymbol{\theta})}{\partial \theta_i \partial \theta_j} \right]$$

where the derivatives are evaluated at the true value of $\boldsymbol{\theta}$ and the expectation is taken with respect to $p(\mathbf{x}; \boldsymbol{\theta})$. Furthermore, an unbiased estimator may be found that attains the bound in that $\mathbf{C}_{\hat{\boldsymbol{\theta}}} = \mathbf{I}^{-1}(\boldsymbol{\theta})$ if and only if

$$\frac{\partial \ln p(\mathbf{x}; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \mathbf{I}(\boldsymbol{\theta})(\mathbf{g}(\mathbf{x}) - \boldsymbol{\theta}) \quad (3.25)$$

for some p -dimensional function $\mathbf{g}$ and some $p \times p$ matrix $\mathbf{I}$. That estimator, which is the MVU estimator, is $\hat{\boldsymbol{\theta}} = \mathbf{g}(\mathbf{x})$, and its covariance matrix is $\mathbf{I}^{-1}(\boldsymbol{\theta})$.

تعمیم به حالت برداری

□ مثال ۸. به دست آوردن تخمینگر MVU برای مثال قبل

$$\frac{\partial \ln p(\mathbf{x}; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \begin{bmatrix} \frac{\partial \ln p(\mathbf{x}; \boldsymbol{\theta})}{\partial A} \\ \frac{\partial \ln p(\mathbf{x}; \boldsymbol{\theta})}{\partial B} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sigma^2} \sum_{n=0}^{N-1} (x[n] - A - Bn) \\ \frac{1}{\sigma^2} \sum_{n=0}^{N-1} (x[n] - A - Bn)n \end{bmatrix}$$

□ رابطه فوق را می توان به صورت زیر بازنویسی نمود:

$$\frac{\partial \ln p(\mathbf{x}; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \begin{bmatrix} \frac{N}{\sigma^2} & \frac{N(N-1)}{2\sigma^2} \\ \frac{N(N-1)}{2\sigma^2} & \frac{N(N-1)(2N-1)}{6\sigma^2} \end{bmatrix} \begin{bmatrix} \hat{A} - A \\ \hat{B} - B \end{bmatrix}$$

تعمیم به حالت برداری

□ ادامه مثال ۸

□ که:

$$\hat{A} = \frac{2(2N-1)}{N(N+1)} \sum_{n=0}^{N-1} x[n] - \frac{6}{N(N+1)} \sum_{n=0}^{N-1} nx[n]$$

$$\hat{B} = -\frac{6}{N(N+1)} \sum_{n=0}^{N-1} x[n] + \frac{12}{N(N^2-1)} \sum_{n=0}^{N-1} nx[n].$$

□ شرط تساوی برقرار است و بنابراین تخمینگرهای فوق MVU هستند.

□ این یک مثال خاص از مدل خطی بود. به طور کلی اگر بتوان داده را به صورت خطی مدل کرد، می توان تخمینگر MVU را به راحتی به دست آورد. (فصل ۴)

تبدیل پارامترها در حالت برداری

□ صورت مسأله:

Assume that it is desired to estimate $\alpha = \mathbf{g}(\boldsymbol{\theta})$ for $\mathbf{g}$, an r -dimensional function.

□ آنگاه:

$$\mathbf{C}_{\hat{\alpha}} - \frac{\partial \mathbf{g}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \mathbf{I}^{-1}(\boldsymbol{\theta}) \frac{\partial \mathbf{g}(\boldsymbol{\theta})^T}{\partial \boldsymbol{\theta}} \geq \mathbf{0}$$

□ ماتریس Jacobian:

$$\frac{\partial \mathbf{g}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \begin{bmatrix} \frac{\partial g_1(\boldsymbol{\theta})}{\partial \theta_1} & \frac{\partial g_1(\boldsymbol{\theta})}{\partial \theta_2} & \cdots & \frac{\partial g_1(\boldsymbol{\theta})}{\partial \theta_p} \\ \frac{\partial g_2(\boldsymbol{\theta})}{\partial \theta_1} & \frac{\partial g_2(\boldsymbol{\theta})}{\partial \theta_2} & \cdots & \frac{\partial g_2(\boldsymbol{\theta})}{\partial \theta_p} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_r(\boldsymbol{\theta})}{\partial \theta_1} & \frac{\partial g_r(\boldsymbol{\theta})}{\partial \theta_2} & \cdots & \frac{\partial g_r(\boldsymbol{\theta})}{\partial \theta_p} \end{bmatrix}.$$

تبدیل پارامترها در حالت برداری

32

□ مثال ۹. تخمین نسبت سیگنال به نویز در اندازه گیری ولتاژ DC

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

□ در اینجا هدف این است که نسبت توان سیگنال به توان نویز (SNR) را تخمین بزنیم.

$$\alpha = \frac{A^2}{\sigma^2}$$

□ قبلا حد کرامر-راو برای تخمین بردار زیر را محاسبه کردیم:

$$\boldsymbol{\theta} = [A \ \sigma^2]^T$$

□ بنابراین برای محاسبه حد تخمین SNR، تابع زیر را در نظر می گیریم:

$$g(\boldsymbol{\theta}) = \theta_1^2 / \theta_2 = A^2 / \sigma^2$$

تبدیل پارامترها در حالت برداری

33

□ ادامه مثال ۹

□ بنابراین مطابق مثال ۶:

$$\mathbf{I}(\boldsymbol{\theta}) = \begin{bmatrix} \frac{N}{\sigma^2} & 0 \\ 0 & \frac{N}{2\sigma^4} \end{bmatrix}.$$

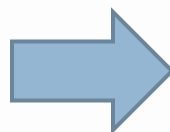
□ محاسبه ماتریس Jacobian:

$$\frac{\partial g(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \begin{bmatrix} \frac{\partial g(\boldsymbol{\theta})}{\partial \theta_1} & \frac{\partial g(\boldsymbol{\theta})}{\partial \theta_2} \end{bmatrix} = \begin{bmatrix} \frac{\partial g(\boldsymbol{\theta})}{\partial A} & \frac{\partial g(\boldsymbol{\theta})}{\partial \sigma^2} \end{bmatrix} = \begin{bmatrix} \frac{2A}{\sigma^2} & -\frac{A^2}{\sigma^4} \end{bmatrix}$$

□ بنابراین:

$$\frac{\partial g(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \mathbf{I}^{-1}(\boldsymbol{\theta}) \frac{\partial g(\boldsymbol{\theta})^T}{\partial \boldsymbol{\theta}} = \begin{bmatrix} \frac{2A}{\sigma^2} & -\frac{A^2}{\sigma^4} \end{bmatrix} \begin{bmatrix} \frac{\sigma^2}{N} & 0 \\ 0 & \frac{2\sigma^4}{N} \end{bmatrix} \begin{bmatrix} \frac{2A}{\sigma^2} \\ -\frac{A^2}{\sigma^4} \end{bmatrix}$$

$$\begin{aligned} &= \frac{4A^2}{N\sigma^2} + \frac{2A^4}{N\sigma^4} \\ &= \frac{4\alpha + 2\alpha^2}{N}. \end{aligned}$$



$$\text{var}(\hat{\alpha}) \geq \frac{4\alpha + 2\alpha^2}{N}.$$

حالت کلی برای مدل گوسی

□ در حالت کلی اگر داده مدل برداری گوسی زیر را داشته باشد:

$$\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}(\boldsymbol{\theta}), \mathbf{C}(\boldsymbol{\theta}))$$

□ آنگاه می توان ماتریس اطلاعات فشر را به صورت زیر محاسبه نمود:

$$[\mathbf{I}(\boldsymbol{\theta})]_{ij} = \left[\frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \theta_i} \right]^T \mathbf{C}^{-1}(\boldsymbol{\theta}) \left[\frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \theta_j} \right] + \frac{1}{2} \text{tr} \left[\mathbf{C}^{-1}(\boldsymbol{\theta}) \frac{\partial \mathbf{C}(\boldsymbol{\theta})}{\partial \theta_i} \mathbf{C}^{-1}(\boldsymbol{\theta}) \frac{\partial \mathbf{C}(\boldsymbol{\theta})}{\partial \theta_j} \right]$$

□ که:

$$\frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \theta_i} = \begin{bmatrix} \frac{\partial [\boldsymbol{\mu}(\boldsymbol{\theta})]_1}{\partial \theta_i} \\ \frac{\partial [\boldsymbol{\mu}(\boldsymbol{\theta})]_2}{\partial \theta_i} \\ \vdots \\ \frac{\partial [\boldsymbol{\mu}(\boldsymbol{\theta})]_N}{\partial \theta_i} \end{bmatrix} \quad \frac{\partial \mathbf{C}(\boldsymbol{\theta})}{\partial \theta_i} = \begin{bmatrix} \frac{\partial [\mathbf{C}(\boldsymbol{\theta})]_{11}}{\partial \theta_i} & \frac{\partial [\mathbf{C}(\boldsymbol{\theta})]_{12}}{\partial \theta_i} & \cdots & \frac{\partial [\mathbf{C}(\boldsymbol{\theta})]_{1N}}{\partial \theta_i} \\ \frac{\partial [\mathbf{C}(\boldsymbol{\theta})]_{21}}{\partial \theta_i} & \frac{\partial [\mathbf{C}(\boldsymbol{\theta})]_{22}}{\partial \theta_i} & \cdots & \frac{\partial [\mathbf{C}(\boldsymbol{\theta})]_{2N}}{\partial \theta_i} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial [\mathbf{C}(\boldsymbol{\theta})]_{N1}}{\partial \theta_i} & \frac{\partial [\mathbf{C}(\boldsymbol{\theta})]_{N2}}{\partial \theta_i} & \cdots & \frac{\partial [\mathbf{C}(\boldsymbol{\theta})]_{NN}}{\partial \theta_i} \end{bmatrix}.$$