باسمه تعالی

دانشگاه علم و صنعت ایران

تئوری تخمین و فیلترهای بهینه

Lecture 3

تخمین بهینه

مقدمه

- تخمین بهینه عبارت است از بهترین حدس ممکن.
- در این بخش پس از ارائه مباحث بنیادی تخمین بهینه و خواص مطلوب مد نظر در آن، سه مبحث مهم تخمین بهینه ارائه می گردد:
  - **Maximum Likelihood**
  - **Maximum a Posteriori**
  - **Minimum Mean-Square Error**

که هر یک از این سه روش منجر به تخمین های متفاوت و تخمینگرهای مختلفی می شوند.

فرموله سازی

- سیگنال $s(t)$ و فرم گسسته آن به صورت $s(n)$ را در نظر بگیرید. همچنین اندازه گیری صورت گرفته $z(n)$ قابل بازنویسی به صورت زیر است:

$$z(n) = g(s(n), v(n), n)$$

که $v(n)$ سیگنال نویز است.

- **هدف**: یافتن تخمین بهینه $s(n)$ است در مرحله n با استفاده از اندازه گیری های $z(1),\ z(2),\ \ldots,\ z(n)$ که می توان آن را به صورت زیر بیان کرد:

$$\boxed{\hat{s}(n) = \alpha_n(z(1),\ z(2),\ \ldots,\ z(n))}$$

فرموله سازی

$$\boxed{\hat{s}(n) = \alpha_n(z(1),\ z(2),\ \ldots,\ z(n))}$$

- پر واضح است که این رابطه نشان می دهد که تخمینگر ممکن است غیرخطی و/یا متغیر با زمان باشد.
- به منظور تعریف یک قضیه بهینه سازی و خروج از ابهامات مساله کلی فوق سیگنال های $s(n)$ و $v(n)$ به صورت سیگنال های تصادفی با توزیع مشترک (**jointly distributed**) روی فضای احتمال S در نظر گرفته می شوند. لذا سیگنال اندازه گیری $z(n)$ نیز یک سیگنال تصادفی روی فضای احتمال S خواهد بود.

$$z(n) = g(s(n), v(n), n)$$

- در نتیجه سیگنال تخمین زده شده $\hat{s}(n)$ نیز یک سیگنال تصادفی بر روی فضای احتمال S به صورت $\hat{s}(n) = \alpha_n(z(1),\ z(2),\ \ldots,\ z(n))$ خواهد بود.

تعریف مساله تخمین بهینه

مطلوب است طراحی تخمینگر بهینه و یافتن تخمین بهینه $s(n)$ از سیگنال $\hat{s}(n)$ با داشتن مقادیر اندازه گیری شده $z(1),\ z(2),\ \ldots,\ z(n)$ ، شناختن تابع g و آشنایی با قضایای بهینه سازی.

$$\boxed{\hat{s}(n)=\alpha_n\left(z(1),\ z(2),\ \ldots,\ z(n)\right)}$$

سه نوع مهم از مسائل تخمین

Filtering: در این حالت سیگنال $s(n)$ در لحظه n با استفاده از اندازه گیری های $z(1),\ z(2),\ \ldots,\ z(n)$ تخمین زده می شود.

Prediction: در این حالت مقدار سیگنال $s(.)$ در لحظات آینده (بعد از n) با استفاده از مشاهدات $z(1),\ z(2),\ \ldots,\ z(n)$ تخمین زده می شود. به عنوان مثال تخمین $s(n+1)$ و $s(n+20)$.

Smoothing: در این حالت مقدار سیگنال $s(.)$ در لحظات قبل از n با استفاده از مشاهدات $z(1),\ z(2),\ \ldots,\ z(n)$ تخمین زده می شود. به عنوان مثال تخمین $s(n-1)$ و $s(n-16)$.

سه نوع مهم از مسائل تخمین

بنابراین می توان مساله تخمین را به صورت کلی زیر بیان نمود که:

"یافتن بهترین حدس یا تخمین ممکن برای مقدار $s(l)$ با استفاده از مشاهدات $z(1), z(2), \ldots, z(n)$."

l	**Estimate**	**Estimation Problem**
n	$\hat{s}(n)$	Filtering
$n+1$	$\hat{s}(n+1)$	One-step prediction
$n+m, m>0$	$\hat{s}(n+m)$	*m*-step prediction
$n-1$	$\hat{s}(n-1)$	Smoothing with lag 1
$n-m, m>0$	$\hat{s}(n-m)$	Smoothing with lag *m*
m constant	$\hat{s}(m)$	Fixed-point smoothing

خواص تخمین

به منظور ارزیابی یک تخمینگر لازم است شرایطی معرفی گردد که بتوان با استفاده از آن رفتار تخمینگر را بررسی نمود.

۱- *Unbiased Estimation*

یکی از مهمترین خواص لازم برای یک تخمین مناسب عبارت است از اینکه:

$$\boxed{E\left[\hat{s}(n)\right]=E\left[s(n)\right]}$$

به عبارت دیگر میانگین خطای تخمین $\tilde{s}(n)$ لازم است برابر صفر باشد

$$E\left[\tilde{s}(n)\right]=0$$

خواص تخمین

برای تخمین هایی که وابسته به ترم های زمانی هستند و تخمیین بدون بایاس به صورت *Asymptotically Unbiased Estimation* مطرح می گردد (که شرط ضعیف تری محسوب می گردد):

$$\lim_{n\to\infty} E\left[\hat{s}(n)\right] = E\left[s(n)\right]$$

۲- *Consistent Estimator*

معیار دیگر که میتوان برای ارزیابی تخمینگر ها مورد استفاده قرار می گیرد، بررسی وضعیت میانگین مربعات خطای(MSE) تخمین است.

اگر   خطای تخمین برابر صفر باشد آنگاه به آن تخمینگر یک تخمینگر consistent گویند.

$$\boxed{E\left[\tilde{s}^2(n)\right] = 0}$$

خواص تخمین

اگر هر دو شرط ۱ و ۲ در مورد یک تخمین برقرار باشد به آن <u>تخمین کامل</u> گویند.

$$\boxed{\left.\begin{array}{l} 1-\begin{bmatrix} E\left[\hat{s}(n)\right] = E\left[s(n)\right] \\ \vee \\ \lim\limits_{n\to\infty} E\left[\hat{s}(n)\right] = E\left[s(n)\right] \end{bmatrix} \\ 2-E\left[\tilde{s}^2(n)\right] = 0 \end{array}\right\} \Rightarrow \hat{s}(n) \text{ is perfect estimate of } s(n)}$$

خواص تخمین

مثال: تحلیل فیلتر میانگین

$s(n)=s$
$z(n)=s+v(n)$
$v(n): \textit{zero mean additive white noise}$
$s, v(n)$ are independent

$$\text{mean filter} \rightarrow \quad \hat{s}(n)=\frac{1}{n}\sum_{j=1}^{n} z(j)$$

$$\begin{aligned} E[\hat{s}(n)] &= \frac{1}{n}\sum_{j=1}^{n} E[z(j)] = \frac{1}{n}\sum_{j=1}^{n} E[s+v(j)] \\ &= \frac{1}{n}\sum_{j=1}^{n}\left(E[s]+E[v(j)]\right) \\ &= \frac{1}{n}\sum_{j=1}^{n}\left(E[s]\right) = E[s] \end{aligned}$$

❋ در نتیجه این تخمینگر یک تخمینگر ***unbiased*** است.

خواص تخمین

ادامه مثال: تحلیل فیلتر میانگین

$$E\left[\tilde{s}^2(n)\right]=E\left[\left(s-\hat{s}(n)\right)^2\right]=E\left[s^2-2s\hat{s}(n)+\hat{s}^2(n)\right]=E\left[s^2\right]-2\underbrace{E\left[s\hat{s}(n)\right]}_{?}+\underbrace{E\left[\hat{s}^2(n)\right]}_{?}$$

$$E\left[s\hat{s}(n)\right]=\frac{1}{n}\sum_{j=1}^{n}E\left[sz(j)\right]=\frac{1}{n}\sum_{j=1}^{n}E\left[s\left(s+v(j)\right)\right]=\frac{1}{n}\sum_{j=1}^{n}E\left[s^2+sv(j)\right]=E\left[s^2\right]$$

$$\begin{aligned} E\left[\hat{s}^2(n)\right] &= \frac{1}{n^2}E\left[\left(z(1)+z(2)+\cdots+z(n)\right)\left(z(1)+z(2)+\cdots+z(n)\right)\right] \\ &= \frac{1}{n^2}E\left[\left(ns+v(1)+v(2)+\cdots+v(n)\right)\left(ns+v(1)+v(2)+\cdots+v(n)\right)\right] \\ &= \frac{1}{n^2}\left\{\left(nE\left[s^2\right]+\sigma_v^2\right)n\right\} = E\left[s^2\right]+\frac{\sigma_v^2}{n} \end{aligned}$$

خواص تخمین

ادامه مثال: تحلیل فیلتر میانگین

$$\Rightarrow\ E\left[\tilde{s}^2(n)\right]=E\left[s^2\right]-2E\left[s^2\right]+E\left[s^2\right]+\frac{\sigma_v^2}{n}=\frac{\sigma_v^2}{n}$$

$$\Rightarrow\ \lim_{n\to\infty} E\left[\tilde{s}^2(n)\right]=\lim_{n\to\infty}\frac{\sigma_v^2}{n}=0$$

※ در نتیجه با افزایش n این تخمینگر یک تخمینگر *Consistent* خواهد بود. و در نتیجه این تخمینگر با شرایط مطرح شده یک تخمینگر کامل است.

تخمین ماکزیمم احتمال

(Maximum Likelihood Estimation)

متغیر تصادفی x با تابع چگالی تصادفی *unimodal* $f_x(x)$ را در نظر بگیرید. پیش از این گفته شد که محتمل ترین مقدار x متناظر است با مقداری از x که موجب ماکزیمم شدن $f_x(x)$ می گردد.

※ با توجه به توضیح فوق و با فرض اینکه مقادیر مشاهده های انجام شده را با می توان به صورت $z=g(s,v)$ بیان نمود، پیشنهاد می گردد که تخمین s با یافتن مقداری از s بدست آید که محتمل ترین مقدار آن برای تولید z است.

$\hat{s}_{ML}$ = values of s that maximizes $f_z(z\,|\,\mathbf{s}=s)$

where, $f_z(z\,|\,\mathbf{s}=s)$: *likelihood function*

تخمین ماکزیمم احتمال

(Maximum Likelihood Estimation)

به عبارت دیگر:

$$\hat{s}_{ML} = \text{values of } s \text{ for which } \frac{\partial f_z(z \mid \mathbf{s}=s)}{\partial s} = 0 \qquad (3.9)$$

همچنین با توجه به اینکه تابع لگاریتم طبیعی یک تابع بطور یکنواخت افزایشی است لذا می توان رابطه (۳.۹)را به صورت زیر بازنویسی نمود و با عنوان *log-likelihood function* شناخته می شود.

$$\hat{s}_{ML} = \text{values of } s \text{ for which } \frac{\partial \ln f_z(z \mid \mathbf{s}=s)}{\partial s} = 0 = \alpha(z) \qquad (3.10)$$

تخمین ماکزیمم احتمال

(Maximum Likelihood Estimation)

مثال(۳.۲): (تخمین **ML**)

$$f_{s,z}(s,z) = \begin{cases} \frac{1}{12}(s+z)e^{-z}, & 0 \le s \le 4,\ \ 0 \le z < \infty \\ 0, & otherwise \end{cases}$$

هدف یافتن تخمین **ML**، s است از مشاهدات z

$$f_z(z \mid \mathbf{s}=s) = \frac{f_{s,z}(s,z)}{f_s(s)}$$

$$f_s(s) = \int_{-\infty}^{\infty} f_{s,z}(s,z)dz = \int_0^{\infty} \frac{1}{12}(s+z)e^{-z}dz = \frac{1}{12}\left[-se^{-z} - ze^{-z} - e^{-z}\right]_0^{\infty} = \frac{s+1}{12}$$

تخمین ماکزیمم احتمال

(Maximum Likelihood Estimation)

ادامه مثال: (تخمین ML)

$$f_z(z\,|\,\mathbf{s}=s)=\frac{s+z}{s+1}e^{-z},\qquad 0\le s\le 4,\quad 0\le z<\infty$$

$$\frac{\partial f_z(z\,|\,\mathbf{s}=s)}{\partial s}=\frac{\partial}{\partial s}\frac{s+z}{s+1}e^{-z}=\frac{1-z}{(s+1)^2}e^{-z}$$

:

$$\begin{cases} z>1, & \dfrac{\partial f_z(z\,|\,\mathbf{s}=s)}{\partial s}<0 \;\Rightarrow\; \{\text{s=0}\rightarrow \max f_z(z\,|\,\mathbf{s}=s)\} \\ 0\le z<1, & \dfrac{\partial f_z(z\,|\,\mathbf{s}=s)}{\partial s}>0 \;\Rightarrow\; \{\text{s=4}\rightarrow \max f_z(z\,|\,\mathbf{s}=s)\} \\ z=1, & \dfrac{\partial f_z(z\,|\,\mathbf{s}=s)}{\partial s}=0 \;\Rightarrow\; \{\text{s=arbitrarily value=}22/9\} \end{cases}$$

تخمین ماکزیمم احتمال

(Maximum Likelihood Estimation)

ادامه مثال: (تخمین ML)

$$\hat{s}_{ML}=\begin{cases} 0, & 1\le z<\infty \\ \dfrac{22}{9}, & z=1 \\ 4, & 0\le z<1 \end{cases}$$

:

تخمین ماکزیمم احتمال

(Maximum Likelihood Estimation)

مثال(۳.۳): (تخمین **ML** با نویز گوسی)

$$z = s + v$$

$$f_v(v) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{v^2}{2\sigma^2}}$$

با فرض استقلال s و v پیش از این نشان داده شده است که

$$f_z(z \mid \mathbf{s} = s) = f_v(v)\vert_{v=z-s} = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(z-s)^2}{2\sigma^2}}$$

$$\Rightarrow \hat{s}_{ML} = z$$

تخمین ماکزیمم احتمال

(Maximum Likelihood Estimation)

مثال(۳.۴): (آشکارسازی سیگنال با استفاده از تخمین **ML**)

$$z = \begin{cases} s+v, & \text{the signal is present} \\ v, & \text{the signal is not present} \end{cases}$$

$$s, v \in S_1$$

با فرض استقلال s و v. هدف آشکاسازی حضور و عدم حضور سیگنال است. به منظور آشکارسازی حضور سیگنال با استفاده از تخمین ML متغیر تصادفی c در فضای احتمال S_2 شامل دو رویداد حضور و/یا عدم حضور سیگنال

$$c = \begin{cases} 1, & \text{the signal is present} \\ 0, & \text{the signal is absent} \end{cases}$$

$$\Rightarrow z = cs + v \qquad \in S_1 \times S_2$$

تخمین ماکزیمم احتمال

(Maximum Likelihood Estimation)

ادامه مثال(۳.۴): (آشکارسازی سیگنال با استفاده از تخمین ML)

$\hat{c}_{ML}$ = values of i that maximizes the likelihood function $f_z(z \mid c = i)$

با توجه به دو مقداره بودن تابع:

$$\Rightarrow \quad \hat{c}_{ML} = \begin{cases} 1, & f_z(z \mid c = 1) > f_z(z \mid c = 0) \\ 0, & \text{otherwise} \end{cases}$$

تخمین ماکزیمم احتمال

(Maximum Likelihood Estimation)

مثال: (تخمین ML از مشاهدات متعدد)

$$s(n) = s$$
$$z(n) = s + v(n)$$
$$v(n) \sim N(0, \sigma^2)$$

با فرض استقلال s و v برای هر مقدار n.

$$Z_n = \begin{bmatrix} z(1) \\ z(2) \\ \vdots \\ z(n) \end{bmatrix}, \quad V_n = \begin{bmatrix} v(1) \\ v(2) \\ \vdots \\ v(n) \end{bmatrix}$$

با توجه به تعریف تخمین ML:

$\hat{s}_{ML}$: value of s that maximizes the likelihood function $f_{Z_n}(Z_n \mid \mathbf{s} = s)$

تخمین ماکزیمم احتمال

(Maximum Likelihood Estimation)

ادامه مثال: (تخمین ML از مشاهدات متعدد)

به منظور محاسبه $f_{Z_n}(Z_n \mid \mathbf{s} = s)$ لازم است $f_{V_n}(V_n)$ محاسبه گردد که با توجه به توصیف $v(n)$ خواهیم داشت:

$$f_{V_n}(V_n) = \frac{1}{(2\pi)^{n/2}\,|P_n|^{1/2}} e^{-\frac{1}{2}V_n^T P_n^{-1} V_n}$$

$$P_n = Cov[V_n]$$

$$|P_n| : \text{determinant of } P_n$$

حال بردارمشاهدات را می توان به صورت زیر بازنویسی نمود:

$$Z_n = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} s + V_n = \beta s + V_n$$

تخمین ماکزیمم احتمال

(Maximum Likelihood Estimation)

ادامه مثال: (تخمین ML از مشاهدات متعدد)

با توجه به آنچه در بخش پیشین ارائه شد:

$$f_{Z_n}(Z_n \mid \mathbf{s} = s) = \frac{1}{(2\pi)^{n/2}\,|P_n|^{1/2}} e^{-\frac{1}{2}(Z_n-\beta s)^T P_n^{-1}(Z_n-\beta s)}$$

$$\frac{\partial f_{Z_n}(Z_n \mid \mathbf{s} = s)}{\partial s} = 0 \rightarrow \boxed{\hat{s}_{ML} = \left[\beta^T P_n^{-1}\beta\right]^{-1} \beta^T P_n^{-1} Z_n}$$

تخمین ماکزیمم احتمال

(Maximum Likelihood Estimation)

ادامه مثال: (تخمین ML از مشاهدات متعدد)

اگر $P_n = \sigma^2 I$ باشد، آنگاه

$$\beta^T P_n^{-1} \beta = \frac{n}{\sigma^2}$$

$$\beta^T P_n^{-1} = \frac{1}{\sigma^2}\begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}$$

$$\Rightarrow \quad \boxed{\hat{s}_{ML} = \frac{1}{n}\sum_{j=1}^{n} z(j)}$$

تخمین ماکزیمم احتمال

(Maximum Likelihood Estimation)

جمع بندی: روش ML به تابع چگالی احتمال $f_s(s)$ وابسته نیست و به عنوان روش *non-Bayesian* شناخته می شود. به عبارت دیگر این روش تخمین به شناخت سیگنال نیاز ندارد. لذا این روش یکی از کلی ترین روش های تخمین بهینه محسوب می گردد. به عبارت دیگر بسیاری دیگر از ضوابط تخمین بهینه را می توان زیر مجموعه روش ML دانست زمانیکه اطلاع و دانش بیشتری در مورد سیگنال در دست است.

* لازم به ذکر است که نقطه ضعف عمده این روش محاسبه *likelihood function* است که به سادگی امکان پذیر نیست.

Maximum *a posteriori* Estimation

با فرض وجود مشاهدات $z = g(s,v)$ ، محتمل ترین مقدار s به منظور رخ دادن مقداری خواهد بود که تابع چگالی شــرطی $z = f_s(s \mid \mathbf{z} = z)$ را ماکزیمم نماید.

$$\hat{s}_{MAP} = \text{values of } s \text{ for which } \frac{\partial f_s(s \mid \mathbf{z} = z)}{\partial s} = 0 \qquad (3.16)$$

با توجه به رابطه **Bayes** و مستقل بودن $f_z(z)$ از مقدار s رابطه (۳.۱۶) را می توان به صورت زیر بازنویسی نمود:

$$\hat{s}_{MAP} = \text{values of } s \text{ for which } \frac{\partial f_z(z \mid \mathbf{s} = s)\, f_s(s)}{\partial s} = 0 \qquad (3.17)$$

Maximum *a posteriori* Estimation

با توجه به اینکه در این روش تخمین، تــابع چگــالی احتمــال $f_s(s)$ بایــد معلوم باشد تفاوت عمده این روش با روش **ML** کاملا مشــخص مــی گردد.

- با توجه به استفاده از رابطه **Bayes** در روش تخمین **MAP**، ایــن روش به عنوان یکی از فرم های **Bayesian Estimation** محسوب می گردد.

Maximum *a posteriori* Estimation

مثال(۳.۶): تخمین MAP با حضور نویز گوسی

$z = s + v$

$v \sim N(0, \sigma_v^2)$

$s \sim N(\eta_s, \sigma_s^2) \quad \rightarrow \quad f_s(s) = \frac{1}{\sqrt{2\pi}\sigma_s} e^{-\frac{(s-\eta_s)^2}{2\sigma_s^2}}$

$$Ex-2.14 \rightarrow \quad f_z(z \mid \mathbf{s} = s) = \frac{1}{\sqrt{2\pi}\sigma_v} e^{-\frac{(z-s)^2}{2\sigma_v^2}}$$

$$\Rightarrow \quad f_z(z \mid \mathbf{s} = s) f_s(s) = \frac{1}{2\pi\sigma_v\sigma_s} e^{-\frac{(z-s)^2}{2\sigma_v^2} - \frac{(s-\eta_s)^2}{2\sigma_s^2}}$$

$$\frac{\partial f_z(z \mid \mathbf{s} = s) f_s(s)}{\partial s} = 0 \Rightarrow \boxed{\hat{s}_{MAP} = \eta_s + \frac{\sigma_s^2}{\sigma_s^2 + \sigma_v^2}(z - \eta_s)}$$

Maximum *a posteriori* Estimation

ادامه مثال(۳.۶):

$$\sigma_v^2 << \sigma_s^2 \Rightarrow \hat{s}_{MAP} = z \overset{Ex-3.3}{=} \hat{s}_{ML}$$

به عبارت دیگر در صورتی که توان نویز در مقابل توان سیگنال کوچک باشد تخمین MAP و ML برابر هستند.

*** از سوی دیگر اگر $\sigma_s^2 \rightarrow \infty$ تابع چگالی سیگنال گوسی به صورت یک تابع چگالی یکنواخت در خواهد آمد. این حالت معادل نداشتن اطلاعات قبلی در مورد s است.**

Maximum *a posteriori* Estimation

مثال(۳.۷): کاربرد تخمین MAP در آشکار سازی سیگنال

$z = cs + v$

$f_c(c)$ is known

$\hat{c}_{MAP}$: values of c that maximise $f_z(z \mid \mathbf{c} = c)\, f_c(c)$

$$\begin{cases} c = 1 \rightarrow probability = p \\ c = 0 \rightarrow probability = 1 - p \end{cases}$$

$$\Rightarrow \quad f_z(z \mid \mathbf{c} = i)\, f_c(i) = \begin{cases} f_z(z \mid \mathbf{c} = 1)\, p & \text{if } c = 1 \\ f_z(z \mid \mathbf{c} = 0)(1 - p) & \text{if } c = 0 \end{cases}$$

$$\Rightarrow \quad \hat{c}_{MAP} = \begin{cases} 1, & \text{if} \quad f_z(z \mid \mathbf{c} = 1)\, p > f_z(z \mid \mathbf{c} = 0)(1 - p) \\ 0, & \text{otherwise} \end{cases}$$

Maximum *a posteriori* Estimation

ادامه مثال(۳.۷): کاربرد تخمین MAP در آشکاری سیگنال

$z = cs + v$

$f_c(c)$ is known

$$\Rightarrow \quad \hat{c}_{MAP} = \begin{cases} 1, & \text{if} \quad \dfrac{f_z(z \mid \mathbf{c} = 1)}{f_z(z \mid \mathbf{c} = 0)} > \dfrac{1 - p}{p} \\ 0, & \text{otherwise} \end{cases}$$

نسبت $\dfrac{f_z(z \mid \mathbf{c} = 1)}{f_z(z \mid \mathbf{c} = 0)}$ را *likelihood ratio* می نامند.

Minimum Mean-Square Error Estimation

با فرض وجود مشاهدات و اندازه گیری ها: $z = g(s,v)$

در این روش خطای تخمین به صورت $\tilde{s}$ تعریف می گردد:

$$\tilde{s} = s - \hat{s}$$

بر اساس آن MSE (میانگین مربعات خطا) عبارت است از:

$$MSE = E\left[(s-\hat{s})^2\right] = E\left[E\left[(s-\hat{s})^2 \mid z\right]\right] = E\left[E\left[\tilde{s}^2 \mid z\right]\right]$$

و تخمین MMSE عبارت است از محاسبه تخمین سیگنال s با استفاده از مینیمم نمودن میانگین مربعات خطای تخمین.

Minimum Mean-Square Error Estimation

قضیه ۳.۱- برای متغیر تصادفی مشاهدات z ، تخمین MMSE سیگنال s عبارت است از امید ریاضی شرطی:

$$\boxed{\hat{s}_{MMSE} = E\left[\mathbf{s} \mid \mathbf{z}\right]}$$

اثبات: اگر s و z دو متغیر تصادفی با توزیع مشترک با تابع چگالی احتمال $f_{s,z}(s,z)$ باشد آنگاه:

$$MSE = E\left[(s-\hat{s})^2\right] = E\left[\left(s-\alpha(z)\right)^2\right]$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left(s-\alpha(z)\right)^2 f_{s,z}(s,z)dsdz$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left(s-\alpha(z)\right)^2 f_s(s \mid z=z) f_z(z)dsdz = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left[s^2 - 2s\alpha(z) + \alpha^2(z)\right] f_s(s \mid z=z)ds f_z(z)dz$$

Minimum Mean-Square Error Estimation

ادامه اثبات قضیه ۳.۱ : با توجه به اینکه لازم است $\alpha(z)$ به عنوان تخمینی از s به ازای مشاهدات z محاسبه گردد و فقط انتگرال داخلی متغیر مورد نظر را در خود جای می دهد پس اگر این انتگرال به ازای این متغیر می نیمم گردد مقدار MSE می نیمم خواهد بود.

(لازم به ذکر است که انتگرال یک PDF نامنفی بوده و $f_s(s\,|\,\mathbf{z}=z)\geq 0$)

$$\rightarrow\ \text{minimize: } \Phi=\int_{-\infty}^{\infty}\left[s^2-2s\alpha(z)+\alpha^2(z)\right]f_s(s\,|\,z=z)ds$$

$$\frac{\partial\Phi}{\partial\alpha(z)}=0\ \Rightarrow\ \int_{-\infty}^{\infty}\left[-2s+2\alpha(z)\right]f_s(s\,|\,z=z)ds=0$$

$$\Rightarrow\ \alpha(z)\underbrace{\int_{-\infty}^{\infty}f_s(s\,|\,z=z)ds}_{1}=\underbrace{\int_{-\infty}^{\infty}sf_s(s\,|\,z=z)ds}_{E[s|z]}\ \Rightarrow\ \boxed{\hat{s}_{MMSE}=\alpha(z)=E\left[s\,|\,z\right]}$$

Minimum Mean-Square Error Estimation

خواص تخمین MMSE:

۱- تخمینگر حاصل یک تخمین منحصر به فرد ارائه می نماید یا به عبارت دیگر مقدار می نیمم مطلق در این تخمین حاصل می شود که علت آن با توجه به فرم مرتبه دوم MSE قابل توجیه است.

۲- تخمین MMSE یک تخمین *unbiased* است. (این خاصیت در حالیکه تعداد مشاهدات نیز محدود باشد برقرار است)

$$E\left[\hat{s}_{MMSE}\right]=E\left[\alpha(z)\right]=E\left[E\left[s\,|\,z\right]\right]=E\left[s\right]$$

۳- تخمین MMSE یک نوع دیگر از مجموعه ***Bayesian Estimation*** است. چرا که با توجه به استفاده از رابطه Bayes این تخمین به داشتن اطلاعاتی از s نیاز دارد.

Minimum Mean-Square Error Estimation

مثال۳.۸: (تخمین MMSE با حضور نویز گوسی)

$z = s + v$

$v \sim N(0, \sigma_v^2)$

we require knowledge of s, so we assume $s \sim N(\bar{s}, \sigma_s^2)$

Also, assume that s and v are *uncorrelated*

$$\Rightarrow \begin{cases} E[z] = E[s] = \bar{s} \\ Var[z] = Var[s] + Var[v] = \sigma_s^2 + \sigma_v^2 \end{cases}$$

$$\Rightarrow z \sim N(\bar{s}, \sigma_s^2 + \sigma_v^2) \quad \Rightarrow \quad f_z(z) = \frac{1}{2\pi\sqrt{\sigma_s^2 + \sigma_v^2}} e^{-\frac{(z-\bar{s})^2}{2(\sigma_s^2+\sigma_v^2)}}$$

Minimum Mean-Square Error Estimation

ادامه مثال۳.۸: (تخمین MMSE با حضور نویز گوسی)

$$\Rightarrow \quad f_s(s \mid \mathbf{z} = z) = \frac{1}{2\pi\sqrt{\frac{\sigma_s^2\sigma_v^2}{\sigma_s^2+\sigma_v^2}}} e^{-\left[-\frac{(z-\bar{s})^2}{2(\sigma_s^2+\sigma_v^2)} + \frac{(z-s)^2}{2\sigma_v^2} + \frac{(s-\bar{s})^2}{2\sigma_s^2}\right]} = \frac{1}{2\pi\sqrt{\frac{\sigma_s^2\sigma_v^2}{\sigma_s^2+\sigma_v^2}}} e^{-\frac{(s-\hat{s}_\eta)^2}{2\frac{\sigma_s^2\sigma_v^2}{\sigma_s^2+\sigma_v^2}}}$$

$$\Rightarrow \quad \hat{s}_{MMSE} = \bar{s} + \frac{\sigma_s^2}{\sigma_s^2 + \sigma_v^2}(z - \bar{s}) = \hat{s}_\eta = \hat{s}_{MAP}$$

به عبارت دیگر اگر s و v دو متغیر تصادفی *uncorrelated* باشند، در ایــن مثال تخمین MMSE از s معادل تخمین MAP است.

Minimum Mean-Square Error Estimation

قضیه ۳.۲: (اصل تعامد) خطای تخمین در تخمین MMSE برهر تـابعی از مشاهدات (z) عمود است.

$$\boxed{E\left[\left(s-E[s\mid z]\right)\gamma(z)\right]=0}$$

اثبات:

$$E\left[\left(s-E[s\mid z]\right)\gamma(z)\right]=E\left[E\left[\left(s-E[s\mid z]\right)\gamma(z)\mid z\right]\right]$$
$$=E\left[E\left[\left(s-E[s\mid z]\right)\mid z\right]\gamma(z)\right]=E\left[\left(E\left[(s)\mid z\right]-E\left[s\mid z\right]\right)\gamma(z)\right]=0$$

Minimum Mean-Square Error Estimation

قضیه ۳.۳: تخمین $\hat{s}=\alpha(z)$ یک تخمین MMSE از s است اگر و فقط اگر خطای تخمین بر هر تابعی از مشاهدات مانند $\gamma(z)$ عمود باشد.

$$\boxed{\hat{s}=\alpha(z):\mathrm{MMSE}\quad\Leftrightarrow\quad E\left[\left(s-\alpha(z)\right)\gamma(z)\right]=0}$$

اثبات : ❊

توجه : قضیه ۳.۲ یکی از خواص تخمین MMSE را مطرح مـی نمایـد در حالیکه قضیه ۳.۳ یک شرط لازم و کافی را ارائه می کند که بر اسـاس آن می توان تخمین های MMSE را شناخت.

Minimum Mean-Square Error Estimation

قضیه ۳.۴: اگر Z نشانگر مجموعه ای محدود از مشاهدات به صورت

$$\boldsymbol{Z}=\{z(n), n_1 \le n \le n_2\}$$

باشد آنگاه تخمین MMSE عبارت خواهد بود از:

$$\boxed{\hat{s}(n)=E\left[s(n)\,|\,\boldsymbol{Z}\right]}$$

اثبات :

$$E\left[\left(s(n)-\alpha(\boldsymbol{Z})\right)^2 \,|\, \boldsymbol{Z}\right]=E\left[\left(s(n)-E[s(n)\,|\,\boldsymbol{Z}]+E[s(n)\,|\,\boldsymbol{Z}]-\alpha(\boldsymbol{Z})\right)^2 \,|\, \boldsymbol{Z}\right]$$

$$=E\left[\left(s(n)-E[s(n)\,|\,\boldsymbol{Z}]\right)^2 \,|\, \boldsymbol{Z}\right]$$

$$+2E\left[\left(s(n)-E[s(n)\,|\,\boldsymbol{Z}]\right)\left(E[s(n)\,|\,\boldsymbol{Z}]-\alpha(\boldsymbol{Z})\right) \,|\, \boldsymbol{Z}\right]$$

$$+E\left[\left(E[s(n)\,|\,\boldsymbol{Z}]-\alpha(\boldsymbol{Z})\right)^2 \,|\, \boldsymbol{Z}\right]$$

Minimum Mean-Square Error Estimation

ادامه اثبات قضیه ۳.۴:

$$E\left[\left(s(n)-E[s(n)\,|\,\boldsymbol{Z}]\right)\left(E[s(n)\,|\,\boldsymbol{Z}]-\alpha(\boldsymbol{Z})\right) \,|\, \boldsymbol{Z}\right]=$$

$$=E\left[\left(s(n)-E[s(n)\,|\,\boldsymbol{Z}]\right) \,|\, \boldsymbol{Z}\right]\left(E[s(n)\,|\,\boldsymbol{Z}]-\alpha(\boldsymbol{Z})\right)$$

$$=\left\{E[s(n)\,|\,\boldsymbol{Z}]-E\left[E[s(n)\,|\,\boldsymbol{Z}]\,|\,\boldsymbol{Z}\right]\right\}\left(E[s(n)\,|\,\boldsymbol{Z}]-\alpha(\boldsymbol{Z})\right)$$

$$=\left\{E[s(n)\,|\,\boldsymbol{Z}]-E[s(n)\,|\,\boldsymbol{Z})]\right\}\left(E[s(n)\,|\,\boldsymbol{Z}]-\alpha(\boldsymbol{Z})\right)$$

$$=0$$

$$\Rightarrow\ E\left[\left(s(n)-\alpha(\boldsymbol{Z})\right)^2 \,|\, \boldsymbol{Z}\right]=E\left[\left(s(n)-E[s(n)\,|\,\boldsymbol{Z}]\right)^2 \,|\, \boldsymbol{Z}\right]+E\left[\left(E[s(n)\,|\,\boldsymbol{Z}]-\alpha(\boldsymbol{Z})\right)^2 \,|\, \boldsymbol{Z}\right]$$

$$\ge E\left[\left(s(n)-E[s(n)\,|\,\boldsymbol{Z}]\right)^2 \,|\, \boldsymbol{Z}\right]$$

$$\overset{E[\ ]}{\Rightarrow}\ \text{MSE}=E\left[\left(s(n)-\alpha(\boldsymbol{Z})\right)^2\right]\ge E\left[\left(s(n)-E[s(n)\,|\,\boldsymbol{Z}]\right)^2\right]$$

$$\Rightarrow\ \boxed{\hat{s}_{MMSE}=E[s(n)\,|\,\boldsymbol{Z}]}$$

Linear MMSE Estimation

با توجه به اینکه محاسبه تخمین MMSE وابسته به محاسبه $f_s(s\,|\,\mathbf{z}=z)$ است و محاسبه این تابع چگالی شرطی میتواند بسیار مشــکل باشــد، لذا به دنبال جایگزینی تابع $\alpha(z)$ با کلاس خاصی از تابع ها هستیم تا با محدودتر نمودن مساله یک حل ممکن ارائه گردد.

کلاس مهم و قدرتمندی از توابع که می تواند به عنوان تخمین بهینــه در نظر گرفته شود، تخمین خطی است که در آن

$$\hat{s}=\alpha(z)=\lambda z$$

محاسبه تخمین LMMSE بسیار ساده تر از MMSE است این در حالیت که عملکرد تخمینگر هم چندان افتی نخواهد داشت.

Linear MMSE Estimation

در تخمین LMMSE هدف تعیین و محاسبه λ است.

یکی از روش های تعیین λ محاسبه مستقیم آن با استفاده از می نــیمم سازی MSE است. بطوریکه

$$\mathrm{MSE}=E\left[(s-\lambda z)^2\right]=E\left[s^2-2\lambda sz+\lambda^2 z^2\right]$$

با گرفتن مشتق جزئی نسبت به λ داریم:

$$\frac{\partial \mathrm{MSE}}{\partial \lambda}=-2E[sz]+2\lambda E\left[z^2\right]=0 \quad \Rightarrow \quad \boxed{\lambda=\frac{E[sz]}{E\left[z^2\right]}}$$

و در نتیجه تخمین LMMSE عبارت خواهد بود از:

$$\boxed{\hat{s}_{LMMSE}=\left(\frac{E[sz]}{E\left[z^2\right]}\right)z \qquad (3.43)}$$

Linear MMSE Estimation

توجه به معادله (۳.۴۳) نشان می دهد که محاسبه تخمین LMMSE نیازی به اطلاع از توابع چگالی احتمال ندارد، بلکه می توان از محاسبه و تخمین گشتاورهای مرتبه دوم $E[sz]$ و $E\left[z^2\right]$ استفاده نمود.

$$E[sz] \approx \frac{1}{M}\sum_{m=1}^{M} s_m z_m$$

$$E\left[z^2\right] \approx \frac{1}{M}\sum_{m=1}^{M} z_m^2$$

Linear MMSE Estimation

قضیه ۳.۵: (اصل تعامد در تخمین LMMSE)

اگر $\alpha(z)$ تخمین LMMSE از s باشد با استفاده از مشاهدات z آنگاه خطای تخمین $s-\alpha(z)$ بر تمامی توابع خطی $\gamma(z)$ عمود است.

$$E\left[\left(s-\alpha(z)\right)\gamma(z)\right]=0$$

اثبات:

$$E\left[\left(s-\alpha(z)\right)\gamma(z)\right]=E\left[\left(s-\frac{E[sz]}{E\left[z^2\right]}z\right)\beta z\right]=\beta E\left[sz-\frac{E[sz]}{E\left[z^2\right]}z^2\right]$$

$$=\beta E[sz]-\beta\frac{E[sz]}{E\left[z^2\right]}E\left[z^2\right]=\beta E[sz]-\beta E[sz]=0$$

Linear MMSE Estimation and Vector RVs

اگر $s_{m\times 1}$ و $z_{q\times 1}$ متغیرهای تصادفی برداری با توزیع مشـــترک باشــند. در خصوص محاسبه تخمین LMMSE به فرم $\hat{s} = Mz$ می توان به صــورت زیر عمل نمود:

$$\text{Let} \quad P = E\left[(s-\hat{s})(s-\hat{s})^T\right]$$

میانگین مربعات خطا را می توان با اســتفاده از تــابع $trace$ مــاتریس P توصیف نمود:

$$MSE = trace[P] = E\left[(s-\hat{s})^T(s-\hat{s})\right] = E\left[(s-Mz)^T(s-Mz)\right]$$

حال به منظور محاسبه ماتریس M میتوان از $trace[P]$ نسبت به ماتریس M مشتق گرفت.

Linear MMSE Estimation and Vector RVs

لذا:

$$\begin{aligned} P &= E\left[(s-Mz)(s-Mz)^T\right] \\ &= E\left[ss^T\right] - E\left[sz^T\right]M^T - M.E\left[zs^T\right] + M.E\left[zz^T\right]M^T \\ \Rightarrow trace(P) &= trace\left(E\left[ss^T\right]\right) - trace\left(E\left[sz^T\right]M^T\right) \\ &\quad - trace\left(ME\left[zs^T\right]\right) + trace\left(M.E\left[zz^T\right]M^T\right) \end{aligned}$$

از طرفی:

$$\begin{cases} \dfrac{\partial trace\left(ABA^T\right)}{\partial A} = 2AB \\ \dfrac{\partial trace(AB)}{\partial A} = \dfrac{\partial trace\left(B^T A^T\right)}{\partial A} = B^T \end{cases}$$

Linear MMSE Estimation and Vector RVs

در نتیجه:

$$\Rightarrow \frac{\partial trace(P)}{\partial M} = -2E\left[sz^T\right] + 2M.E\left[zz^T\right] = 0 \qquad (3.50)$$

$$\Rightarrow \boxed{M = E\left[sz^T\right]\left(E\left[zz^T\right]\right)^{-1}} \qquad (3.51)$$

$$\Rightarrow \hat{s} = E\left[sz^T\right]\left(E\left[zz^T\right]\right)^{-1} z \qquad (3.52)$$

$$\begin{cases} \dfrac{\partial trace\left(ABA^T\right)}{\partial A} = 2AB \\ \dfrac{\partial trace\left(AB\right)}{\partial A} = \dfrac{\partial trace\left(B^T A^T\right)}{\partial A} = B^T \end{cases}$$

Linear MMSE Estimation

قضیه ۳.۶: (اصل تعامد در تخمین LMMSE و متغیرهای تصادفی برداری)

اگر $s_{m\times1}$ و $z_{q\times1}$ متغیرهای تصادفی با توزیع مشترک باشند و تخمین LMMSE از s با استفاده از مشاهدات z باشد، آنگاه خطای تخمین بر z عمود است.

$$E\left[\left(s-\hat{s}\right)z^T\right] = 0_{m\times q}$$

اثبات:

$$E\left[\left(s-\hat{s}\right)z^T\right] = E\left[\left\{s - E\left[sz^T\right]\left(E\left[zz^T\right]\right)^{-1} z\right\} z^T\right]$$

$$= E\left[sz^T\right] - E\left[E\left[sz^T\right]\left(E\left[zz^T\right]\right)^{-1} zz^T\right]$$

$$= E\left[sz^T\right] - E\left[sz^T\right]\left(E\left[zz^T\right]\right)^{-1} E\left[zz^T\right] = 0$$

Linear MMSE Estimation

قضیه ۳.۷:

اگر $\alpha(z)$ یک تخمین خطی از s با استفاده از مشاهدات z باشد، آنگاه α مقدار MSE را می نیمم می نماید اگر و فقط اگر خطای تخمین $s-\alpha(z)$ بر مشاهدات z عمود باشد.

$$E\left[\left(s-\alpha(z)\right)z^T\right]=0_{m\times q}$$

توجه: با استفاده از این قضیه میتوان تخمینگر خطی نامتغیر بازمان MMSE را ساده تر طراحی نمود.

Linear MMSE Estimation

به عنوان مثال در حالت اسکالر

$$E\left[\left(s-\lambda z\right)\beta z\right]=0$$

$$\Rightarrow \quad \beta E[sz]=\beta\lambda E\left[z^2\right] \quad \Rightarrow \quad \lambda=\frac{E[sz]}{E\left[z^2\right]}$$

با فرض اینکه میانگین s و z برابر صفر باشند

$$\lambda=\frac{E[sz]-E[s]E[z]}{E\left[z^2\right]-\left(E[z]\right)^2}=\frac{Cov[s,z]}{\sigma_z^2}$$

با تعریف *Correlation Coefficient* به صورت:

$$\rho_{s,z}=\frac{Cov[s,z]}{\sqrt{\sigma_s^2\sigma_z^2}}$$

Linear MMSE Estimation

$$\lambda = \frac{E[sz]-E[s]E[z]}{E[z^2]-(E[z])^2} = \frac{Cov[s,z]}{\sigma_z^2} = \rho_{s,z}\frac{\sigma_s}{\sigma_z}$$

$$\boxed{\Rightarrow \quad \hat{s}_{LMMSE} = \rho_{s,z}\frac{\sigma_s}{\sigma_z} z \qquad (3.55)}$$

ضریب وزنی در رابطه (۳.۵۵) قصد دارد موازنه ای ما بین سیگنال و نویز برقرار نماید.

σ_s^2 معیاری از توان سیگنال در مشاهدات **z** را نشان می دهد، بنابراین بزرگ بودن σ_s^2 وزن مشاهدات **z** را افزایش می دهد.

از طرفی σ_z^2 واریانس مشاهدات یا به عبارتی توان نویز را نشان می دهد که در صورتی که توان نویز بزرگ باشد مشاهدات **z** برای تعیین **s** قابل اعتماد نخواهند بود و ضریب وزنی کاهش می یابد.

Linear MMSE Estimation

معادله (۳.۵۵) نشان می دهد که اگر **s** و **z** دو متغیر تصادفی *uncorrelated* باشند و/یا به عبارت دیگر اگر سیگنال و مشاهدات از طریق گشتاورهای مرتبه دو مشابه و مرتبط نباشند آنگاه

$$\rho_{s,z} = 0 \qquad \Rightarrow \quad \hat{s}_{LMMSE} = 0$$

Linear MMSE Estimation

تحلیل عملکرد تخمین LMMSE:

با فرض اینکه میانگین s و z برابر صفر باشند

$$E\left[s-\hat{s}_{LMMSE}\right]=E\left[s-\lambda z\right]=0 \quad \Rightarrow \quad \text{unbiased estimation}$$

$$E\left[\left(s-\hat{s}_{LMMSE}\right)^2\right]=E\left[s^2\right]-2\frac{\sigma_s}{\sigma_z}\rho_{s,z}E\left[sz\right]+\frac{\sigma_s^2}{\sigma_z^2}\rho_{s,z}^2E\left[z^2\right]$$

$$=E\left[s^2\right]-2\frac{\sigma_s}{\sigma_z}\rho_{s,z}Cov(s,z)+\frac{\sigma_s^2}{\sigma_z^2}\rho_{s,z}^2\sigma_z^2=\sigma_s^2-2\frac{\sigma_s}{\sigma_z}\rho_{s,z}\left(\rho_{s,z}\sigma_s\sigma_z\right)+\frac{\sigma_s^2}{\sigma_z^2}\rho_{s,z}^2\sigma_z^2$$

$$=\sigma_s^2\left(1-\rho_{s,z}^2\right)$$

$$\Rightarrow\begin{cases} If \quad \rho_{s,z}=0 \quad \Rightarrow \quad MSE \text{ will be maximaized} \\ If \quad \rho_{s,z}=\pm 1 \quad \Rightarrow \quad MSE=0 \quad \Rightarrow \quad z=\mu s \end{cases}$$

Linear MMSE Estimation

جمع بندی روش MMSE

تخمین MMSE به محاسبه امید ریاضی شرطی $E[s|z]$ بر می گردد و با توجه به مشکل بودن محاسبه آن مساله به یک مساله زیربهینه به نام تخمین LMMSE تبدیل گشت.

توجه: اگر s و z متغیرهای گوسی با توزیع مشترک باشند تخمین LMMSE همان تخمین بهینه MMSE خواهد بود.

مقایسه روشهای تخمین بهینه مطرح شده

روشهای تخمین **ML** و **MAP** به نظر مشابه می آیند ولیکن تفاوت های بسیاری ما بین آنها وجود دارد:

تخمینگر **ML** : با داشتن مشاهدات z، مقداری از s که احتمـال بـالاتری دارد که مشاهدات z را تولید می کنند.

$$\hat{s}_{ML} = \text{values of } s \text{ for which } \frac{\partial f_z(z \mid \mathbf{s} = s)}{\partial s} = 0 \qquad (3.9)$$

تخمینگر **MAP** : با داشتن مشاهدات z، مقداری از s که احتمال بـالاتری برای رخ دادن دارد(شکلی از تخمین ***Bayesian***).

$$\hat{s}_{MAP} = \text{values of } s \text{ for which } \frac{\partial f_s(s \mid \mathbf{z} = z)}{\partial s} = 0 \qquad (3.16)$$

مقایسه روشهای تخمین بهینه مطرح شده

توجه: تخمین **ML** و **MAP** با تعداد کم نمونه ها ممکن است پاسخ بهتری از تخمین **MMSE** در اختیار قرار دهند ولیکن در تعداد زیـاد تخمـین **MMSE** مقدار **MSE** را کمتر خواهد نمود.

توجه: تخمین **LMMSE** عبارت است از مصالحه ای ما بین بهینـه سـازی کامل و ممکن بودن یافتن تخمین. این روش بـه محاسـبه مقـادیر بـا بیشترین احتمال (**likely value**) نیازی ندارد. بلکه جهـت محاسـبه این تخمین لازم است گشتاورهای مرتبه ۲ از s و z محاسبه گردند.لذا نیازی به محاسبه مستقیم مقدار تابع چگالی احتمال نیست. کـه ایـن موجب ساده تر شدن استفاده از ایـن تخمـین در محاسـبه و کـاربرد می گردد.

مقایسه روشهای تخمین بهینه مطرح شده

مثال ۳.۱۰: (تخمین MMSE و ML)

متغیر تصادفی مثال ۳.۲ را درنظر بگیرید.

$$f_{s,z}(s,z)=\begin{cases}\frac{1}{12}(s+z)e^{-z}, & 0\le s\le 4,\ \ 0\le z<\infty\\ 0, & otherwise\end{cases}$$

$$\Rightarrow \quad f_z(z)=\int_0^4 f_{s,z}(s,z)ds=\frac{(z+2)}{3}e^{-z} \quad \Rightarrow \quad f_z(s\,|\,\mathbf{z}=z)=\frac{1}{4}\frac{s+z}{z+2}, \quad 0\le s\le 4, 0\le z<\infty$$

$$\Rightarrow \quad \text{MSE}(\hat{s}_{ML})=\int_0^\infty\int_0^4\left(s-\hat{s}_{ML}\right)^2 f_{s,z}(s,z)dsdz$$

$$=\int_0^1\int_0^4\left(s-4\right)^2\frac{s+z}{12}e^{-z}dsdz+\int_1^\infty\int_0^4\left(s-0\right)^2\frac{s+z}{12}e^{-z}dsdz\simeq 4.8636$$

مقایسه روشهای تخمین بهینه مطرح شده

ادامه مثال ۳.۱۰: (تخمین MMSE و ML)

$$\hat{s}_{MMSE}=E\left[s\,|\,z\right]=\int_0^4 sf_s(s\,|\,z=z)ds=\frac{6z+16}{3z+6}$$

$$\Rightarrow \quad \text{MSE}(\hat{s}_{MMSE})=\int_0^\infty\int_0^4\left(s-\hat{s}_{MMSE}\right)^2 f_{s,z}(s,z)dsdz\simeq 1.1192$$

همانگونه که انتظار می رفت:

- تخمین ML و MMSE دو مقدار متفاوت خواهند داشت.
- تخمین MMSE دارای مقدار MSE کوچتری نسبت به تخمین ML است.