



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

تئوری تخمین

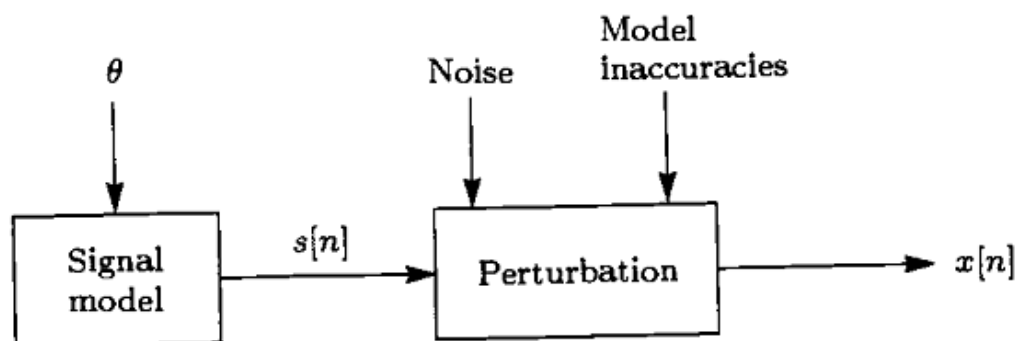
ESTIMATION THEORY

نیم سال اول سال تحصیلی ۰۴-۰۵

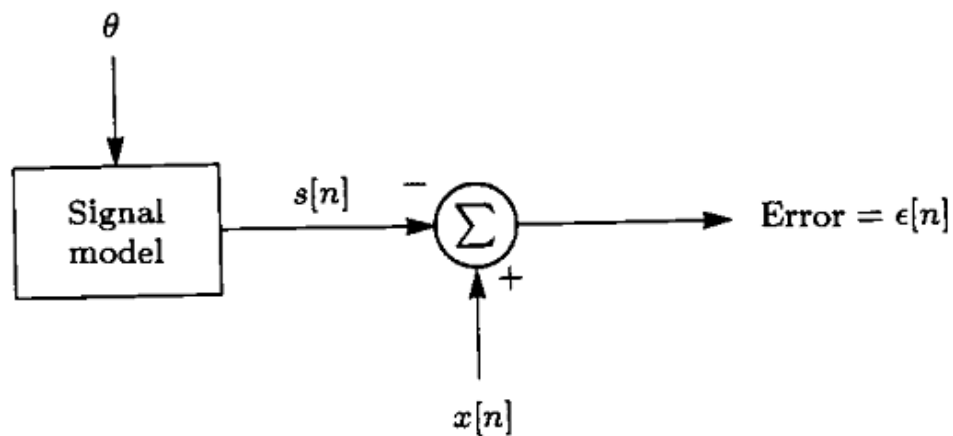
تخمین کمترین مربعات

- مقدمه
- تعریف تخمین LS
- تخمین LS خطی
- تخمین LS وزن دار
- تعبیر هندسی تخمین LS
- تخمین LS مقید
- تخمین LS غیرخطی

- در فصل های قبل به دنبال تخمین گر های بهینه بودیم.
- در این فصل تخمینگری معرفی می شود که هیچ معیار بهینگی ندارد.
- به دلیل در نظر گرفتن معیار کمترین مربعات، این تخمینگر در بسیاری از مسائل نتایج مطلوبی دارد و در عمل استفاده گسترده ای دارد.
- در این تخمینگر نیازی به دانستن هیچ آمارگانی از داده نیست و تنها یک مدل داده در نظر گرفته می شود.
- نقطه ضعف این تخمینگر این است که ادعایی درباره بهینه بودن آن نمی توان داشت و واریانس تخمین جز با فرضیاتی درباره آمارگان داده به دست نمی آید.



(a) Data model



(b) Least squares error

تعریف تخمین LS

5

□ در تخمین LS به دنبال θ ای هستیم که معیار زیر را کمینه می کند:

$$J(\theta) = \sum_{n=0}^{N-1} (x[n] - s[n])^2$$

□ یعنی باید θ ای را پیدا کنیم که داده کمترین فاصله را با سیگنال بدون نویز پیدا کند.

□ هیچ فرض آماری روی داده در نظر گرفته نمی شود.

□ این تخمین معمولاً جایی به کار می رود که اطلاعات آماری از داده نداریم یا اینکه پیدا کردن تخمین های بهینه ناممکن یا خیلی مشکل است.

□ مثال ۱. ولتاژ DC در نویز

□ $X[n] = A + w[n]$

□ در اینجا سیگنال بدون خطا همان ولتاژ است: $s[n] = A$ ، بنابراین تابعی که باید کمینه شود:

$$J(A) = \sum_{n=0}^{N-1} (x[n] - A)^2. \quad \Rightarrow \quad \hat{A} = \frac{1}{N} \sum_{n=0}^{N-1} x[n] = \bar{x}$$

□ آیا می توان ادعا کرد که این تخمین همان MVU است؟

□ در نظر گرفتن مدل اشتباه برای سیگنال می تواند به بایاس تخمینگر منجر شود.

□ مثال ۲

$s[n] = A \cos 2\pi f_0 n$, where f_0 is known and A is to be estimated.

□ تابعی که باید کمینه شود:

$$J(A) = \sum_{n=0}^{N-1} (x[n] - A \cos 2\pi f_0 n)^2$$

$$\frac{\partial J}{\partial A} = -2 \sum_{n=0}^{N-1} (x[n] - A \cos 2\pi f_0 n) \cos 2\pi f_0 n = 0 \Rightarrow A = \frac{\sum_{n=0}^{N-1} x[n] \cos 2\pi f_0 n}{\sum_{n=0}^{N-1} \cos^2 2\pi f_0 n}$$

□ حال فرض کنید مقدار A معلوم است و می خواهیم فرکانس را تخمین بزنیم:

$$J(f_0) = \sum_{n=0}^{N-1} (x[n] - \cos 2\pi f_0 n)^2.$$

□ مثال ۲. ادامه

□ در اینجا به دلیل غیر خطی بودن تابع، به دست آوردن جواب صریح برای فرکانس ممکن نیست و برای حل آن باید از روش های عددی کمک گرفت.

□ این یک مثال از تخمین غیر خطی کمترین مربعات (NLLS) است.

□ در نهایت فرض کنید هر دو پارامتر دامنه و فرکانس مجهول هستند:

$$J(A, f_0) = \sum_{n=0}^{N-1} (x[n] - A \cos 2\pi f_0 n)^2$$

□ در اینجا می توان ابتدا نسبت به A مشتق گرفت و جواب را در معادله قرار داده و سپس از روش عددی نسبت به f_0 حل کرد. (مساله قابل جداسازی است: separable)

□ در اینجا سیگنال بر حسب پارامتر مجهول، خطی است:

$$s[n] = \theta h[n]$$

□ محاسبه تخمین با کمینه کردن تابع زیر:

$$J(\theta) = \sum_{n=0}^{N-1} (x[n] - \theta h[n])^2.$$

مشتق گیری و
مساوی صفر

$$\hat{\theta} = \frac{\sum_{n=0}^{N-1} x[n]h[n]}{\sum_{n=0}^{N-1} h^2[n]}.$$

□ این روند شبیه روند مشاهده شده در فیلتر منطبق است.

□ محاسبه خطای کمترین مربعات:

$$\begin{aligned}
 J_{\min} = J(\hat{\theta}) &= \sum_{n=0}^{N-1} (x[n] - \hat{\theta}h[n])(x[n] - \hat{\theta}h[n]) \\
 &= \sum_{n=0}^{N-1} x[n](x[n] - \hat{\theta}h[n]) - \underbrace{\hat{\theta} \sum_{n=0}^{N-1} h[n](x[n] - \hat{\theta}h[n])}_S \\
 &= \sum_{n=0}^{N-1} x^2[n] - \hat{\theta} \sum_{n=0}^{N-1} x[n]h[n]. \\
 J_{\min} &= \sum_{n=0}^{N-1} x^2[n] - \frac{\left(\sum_{n=0}^{N-1} x[n]h[n] \right)^2}{\sum_{n=0}^{N-1} h^2[n]}.
 \end{aligned}$$

□ می توان نشان داد:

$$0 \leq J_{\min} \leq \sum_{n=0}^{N-1} x^2[n].$$

□ در حالت برداری در مدل خطی تخمین LS داریم: $\mathbf{s} = \mathbf{H}\boldsymbol{\theta}$

□ که $\boldsymbol{\theta}$ یک بردار p تایی و $\mathbf{s}$ یک بردار N تایی است.

□ در این حالت داریم:

$$\begin{aligned} J(\boldsymbol{\theta}) &= \sum_{n=0}^{N-1} (x[n] - s[n])^2 \\ &= (\mathbf{x} - \mathbf{H}\boldsymbol{\theta})^T (\mathbf{x} - \mathbf{H}\boldsymbol{\theta}). \end{aligned}$$

$$\begin{aligned} J(\boldsymbol{\theta}) &= \mathbf{x}^T \mathbf{x} - \mathbf{x}^T \mathbf{H}\boldsymbol{\theta} - \boldsymbol{\theta}^T \mathbf{H}^T \mathbf{x} + \boldsymbol{\theta}^T \mathbf{H}^T \mathbf{H}\boldsymbol{\theta} \\ &= \mathbf{x}^T \mathbf{x} - 2\mathbf{x}^T \mathbf{H}\boldsymbol{\theta} + \boldsymbol{\theta}^T \mathbf{H}^T \mathbf{H}\boldsymbol{\theta} \end{aligned}$$

□ برای کمینه کردن J مشتق می گیریم:

$$\frac{\partial J(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = -2\mathbf{H}^T \mathbf{x} + 2\mathbf{H}^T \mathbf{H} \boldsymbol{\theta}. \quad \Rightarrow \quad \hat{\boldsymbol{\theta}} = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{x}.$$

□ باید دقت کنیم که این تخمین همان BLUE یا MVUE نیست، مگر اینکه شرایط آماری لازم را داشته باشد.

□ محاسبه خطای کمینه:

$$\begin{aligned} J_{\min} &= J(\hat{\boldsymbol{\theta}}) \\ &= (\mathbf{x} - \mathbf{H}\hat{\boldsymbol{\theta}})^T (\mathbf{x} - \mathbf{H}\hat{\boldsymbol{\theta}}) \\ &= (\mathbf{x} - \mathbf{H}(\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{x})^T (\mathbf{x} - \mathbf{H}(\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{x}) \\ &= \mathbf{x}^T (\mathbf{I} - \mathbf{H}(\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T) (\mathbf{I} - \mathbf{H}(\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T) \mathbf{x} \\ &= \mathbf{x}^T (\mathbf{I} - \mathbf{H}(\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T) \mathbf{x}. \end{aligned}$$

$$\begin{aligned} J_{\min} &= \mathbf{x}^T \mathbf{x} - \mathbf{x}^T \mathbf{H} (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{x} \\ &= \mathbf{x}^T (\mathbf{x} - \mathbf{H} \hat{\boldsymbol{\theta}}). \end{aligned}$$

□ حالت کلی تری از تخمین خطی LS، تخمین وزدن دار LS است:
Weighted Least Square

$$J(\boldsymbol{\theta}) = (\mathbf{x} - \mathbf{H}\boldsymbol{\theta})^T \mathbf{W} (\mathbf{x} - \mathbf{H}\boldsymbol{\theta}).$$

□ با استفاده از $\mathbf{W}$ می توان به داده های مطمئن تر، وزن بیشتری داد.

□ مثلاً در مثال تخمین ولتاژ DC، می توان تابع مجموع مربعات را به صورت زیر تعریف کرد:

$$J(A) = \sum_{n=0}^{N-1} w_n (x[n] - A)^2.$$

تخمین LS وزن دار

14

□ در این حالت، ماتریس $\mathbf{W}$ یک ماتریس قطری است که قطر اصلی آن همان وزن های w_n است:

$$J(\mathbf{A}) = (\mathbf{x} - \mathbf{1}\mathbf{A})^T \begin{bmatrix} w_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & w_n \end{bmatrix} (\mathbf{x} - \mathbf{1}\mathbf{A})$$

□ مثلاً اگر وزن هر داده را معکوس واریانس نویز قرار دهیم: $w_n = 1/\sigma_n^2$

□ پس از محاسبه تخمین خواهیم داشت:

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

□ در حالت کلی تخمین WLS به صورت زیر است:

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

$$J_{\min} = \mathbf{x}^T (\mathbf{W} - \mathbf{W} \mathbf{H} (\mathbf{H}^T \mathbf{W} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{W}) \mathbf{x}$$

- از منظر هندسی، هر ماتریس ستونی با N درایه، یک نقطه (بردار) در فضای N بعدی است.
- ماتریس H دارای p ستون است که هر ستون یک بردار در فضای N بعدی محسوب می شود.
- بنابراین در مدل سیگنال خطی داریم:

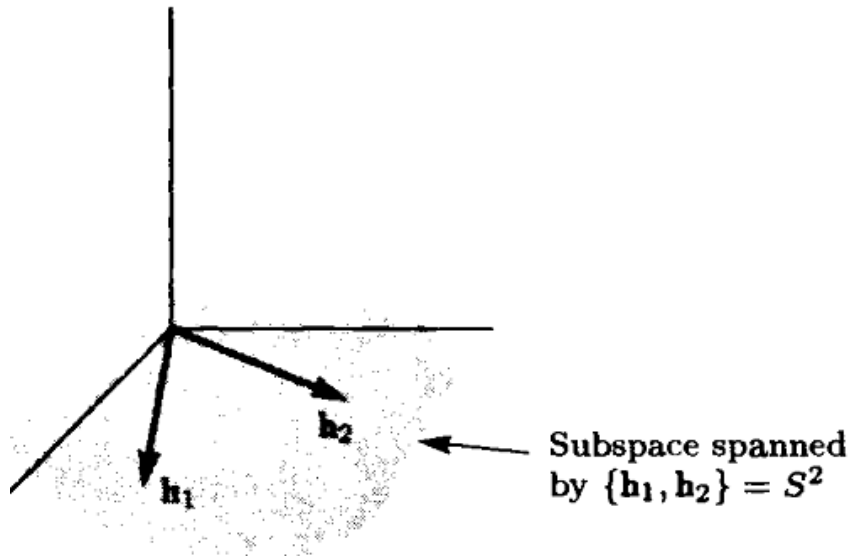
$$\mathbf{s} = \begin{bmatrix} \mathbf{h}_1 & \mathbf{h}_2 & \dots & \mathbf{h}_p \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_p \end{bmatrix}$$
$$= \sum_{i=1}^p \theta_i \mathbf{h}_i$$

- یعنی هر سیگنال به صورت $\mathbf{s} = \mathbf{H}\boldsymbol{\theta}$ یک ترکیب خطی از بردارهای واقع در ستون های ماتریس $\mathbf{H}$ است.
- اگر همه $\boldsymbol{\theta}$ های ممکن را در نظر بگیریم، یک زیرفضا از فضای N بعدی به وسیله ستون های ماتریس $\mathbf{H}$ پوشش داده می شود.
- ضمناً مقدار تابع جمع مربعات، فاصله اقلیدسی بین بردار $\mathbf{H}\boldsymbol{\theta}$ (واقع در زیرفضای $\mathbf{H}$) و بردار داده $\mathbf{x}$ است:

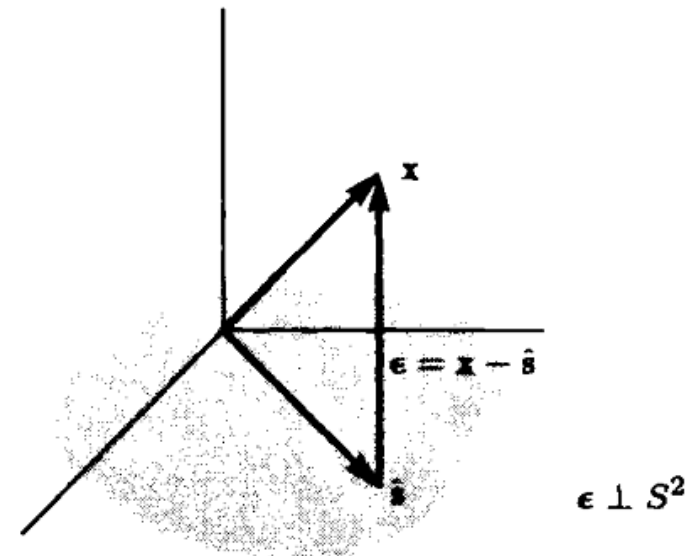
$$J(\boldsymbol{\theta}) = (\mathbf{x} - \mathbf{H}\boldsymbol{\theta})^T (\mathbf{x} - \mathbf{H}\boldsymbol{\theta}).$$

$$\begin{aligned} J(\boldsymbol{\theta}) &= \|\mathbf{x} - \mathbf{H}\boldsymbol{\theta}\|^2 \\ &= \left\| \mathbf{x} - \sum_{i=1}^p \theta_i \mathbf{h}_i \right\|^2. \end{aligned}$$

□ بنابراین در تخمین LS هدف این است که بردار θ ای پیدا کنیم که فاصله $H\theta$ (واقع در زیرفضای H) از x کمینه شود.



(a) Signal subspace



(b) Orthogonal projection to determine signal estimate

□ برای یافتن نزدیک ترین فاصله $\mathbf{x}$ از زیرفضای $\mathbf{H}$ ، باید تصویر (projection) بردار $\mathbf{x}$ در زیرفضای $\mathbf{H}$ پیدا شود. این تصویر، همان تخمین سیگنال یعنی $\hat{\mathbf{S}} = \mathbf{H}\hat{\boldsymbol{\theta}}$ خواهد بود.

□ در این حالت بردار خطا بر تمام بردارهای زیرفضا عمود است. مثلاً اگر $p=2$ باشد داریم:

$$(\mathbf{x} - \mathbf{H}\boldsymbol{\theta})^T [\mathbf{h}_1 \quad \mathbf{h}_2] = \mathbf{0}^T$$

$$(\mathbf{x} - \mathbf{H}\boldsymbol{\theta})^T \mathbf{H} = \mathbf{0}^T.$$

□ و بنابراین:

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

The error vector must be orthogonal to the columns of $\mathbf{H}$.

□ مقدار خطای کمینه، فاصله $\mathbf{x}$ از تصویر خود در زیرفضای $\mathbf{H}$ است:

the minimum LS error is $\|\mathbf{x} - \hat{\mathbf{s}}\|^2$

$$\begin{aligned}\|\mathbf{x} - \hat{\mathbf{s}}\|^2 &= \|\mathbf{x} - \mathbf{H}\hat{\boldsymbol{\theta}}\|^2 \\ &= (\mathbf{x} - \mathbf{H}\hat{\boldsymbol{\theta}})^T (\mathbf{x} - \mathbf{H}\hat{\boldsymbol{\theta}}).\end{aligned}$$

□ هر چه مدل دقیق تر باشد، برداری که در زیرفضای $\mathbf{H}$ شکل می گیرد، به $\mathbf{x}$ نزدیکتر خواهد بود.

□ در حالتی که ستونهای $\mathbf{H}$ (بردارهای پایه زیرفضا) یکه و بر هم عمود باشند، آنگاه تخمین سیگنال همان ترکیب خطی بردارهای پایه با وزن طول بردار $\mathbf{x}$ در آن راستا خواهد بود. (در این حالت داریم: $\mathbf{H}^T \mathbf{H} = \mathbf{I}$)

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

□ گاهی قید هایی (اطلاعات پیشینی) از پارامتر مورد تخمین وجود دارد که می خواهیم تخمین با لحاظ کردن آن قیدها انجام شود.

□ قیدهای خطی روی پارامترها را می توان به صورت زیر بیان کرد:

$$\mathbf{A}\boldsymbol{\theta} = \mathbf{b}$$

□ در این حالت می خواهیم مجموع مربعات با توجه به قید فوق کمینه شود.

□ تابع جدیدی که باید کمینه شود:

$$J_c = (\mathbf{x} - \mathbf{H}\boldsymbol{\theta})^T (\mathbf{x} - \mathbf{H}\boldsymbol{\theta}) + \boldsymbol{\lambda}^T (\mathbf{A}\boldsymbol{\theta} - \mathbf{b})$$

$$J_c = \mathbf{x}^T \mathbf{x} - 2\boldsymbol{\theta}^T \mathbf{H}^T \mathbf{x} + \boldsymbol{\theta}^T \mathbf{H}^T \mathbf{H} \boldsymbol{\theta} + \boldsymbol{\lambda}^T \mathbf{A} \boldsymbol{\theta} - \boldsymbol{\lambda}^T \mathbf{b}.$$

$$\frac{\partial J_c}{\partial \theta} = -2\mathbf{H}^T \mathbf{x} + 2\mathbf{H}^T \mathbf{H} \theta + \mathbf{A}^T \lambda$$

$$\begin{aligned} \hat{\theta}_c &= (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{x} - \frac{1}{2} (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{A}^T \lambda \\ &= \hat{\theta} - (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{A}^T \frac{\lambda}{2} \end{aligned}$$

□ جایگذاری جواب در قید و پیدا کردن مقدار λ :

$$\mathbf{A} \theta_c = \mathbf{A} \hat{\theta} - \mathbf{A} (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{A}^T \frac{\lambda}{2} = \mathbf{b} \quad \Rightarrow \quad \frac{\lambda}{2} = [\mathbf{A} (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{A}^T]^{-1} (\mathbf{A} \hat{\theta} - \mathbf{b}).$$

□ در نهایت:

$$\hat{\theta}_c = \hat{\theta} - (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{A}^T [\mathbf{A} (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{A}^T]^{-1} (\mathbf{A} \hat{\theta} - \mathbf{b})$$

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

□ مثال ۳. فرض کنید سیگنال دریافتی به صورت زیر است:

$$s[n] = \begin{cases} \theta_1 & n=0 \\ \theta_2 & n=1 \\ 0 & n=2 \end{cases}$$

□ و ما داده $\{x[0], x[1], x[2]\}$ را دریافت کرده ایم. تخمین پارامترهای θ_1 و θ_2 را از روش LS حساب کنید.

□ مدل داده به صورت روبرو است: $X[n] = s[n] + w[n]$

□ و ضمناً سیگنال $s[n]$ بر حسب پارامترهای مجهول خطی است:

$$s = H\theta = H \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} \Rightarrow \begin{bmatrix} \theta_1 \\ \theta_2 \\ 0 \end{bmatrix} = \begin{bmatrix} ? & ? \\ ? & ? \\ ? & ? \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} \Rightarrow H = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

□ مثال ۳. ادامه

□ تخمین پارامتر:

$$\hat{\theta} = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{x} = \begin{bmatrix} x[0] \\ x[1] \end{bmatrix}$$

□ حال فرض کنید که از قبل می دانیم: $\theta_1 = \theta_2$. در این حالت قیدی به صورت زیر داریم:

$$\begin{bmatrix} 1 & -1 \end{bmatrix} \theta = 0$$

□ یعنی $\mathbf{A} = [1, -1]$ و $\mathbf{b} = 0$. ضمناً از آنجایی که $\mathbf{H}^T \mathbf{H} = \mathbf{I}$ خواهیم داشت:

$$\hat{\theta}_c = \hat{\theta} - \mathbf{A}^T (\mathbf{A} \mathbf{A}^T)^{-1} \mathbf{A} \hat{\theta} = [\mathbf{I} - \mathbf{A}^T (\mathbf{A} \mathbf{A}^T)^{-1} \mathbf{A}] \hat{\theta}.$$

□ و در نهایت:

$$\hat{\theta}_c = \begin{bmatrix} \frac{1}{2}(x[0] + x[1]) \\ \frac{1}{2}(x[0] + x[1]) \end{bmatrix}$$

تخمین LS غیر خطی

□ در بسیاری از مسائل، مدل سیگنال بر حسب پارامتر مجهول خطی نیست و لذا نمی توان رابطه مجموع مربعات را به راحتی کمینه کرد:

$$J = (\mathbf{x} - s(\boldsymbol{\theta}))^T (\mathbf{x} - s(\boldsymbol{\theta}))$$

□ در اینجا با یک مساله بهینه سازی روبرو هستیم که با توجه به شکل مساله باید از یکی از روش های تحلیلی یا عددی برای حل آن استفاده نمود، مانند روش های جستجوی شبکه ای، روش نیوتن-رافسون و ...

□ یکی از روش های حل مساله در این حالت، تبدیل پارامترها به صورتی است که مساله به فرم خطی تبدیل شود:

تخمین LS غیر خطی

25

$$\alpha = g(\theta)$$

□

$$s(\theta(\alpha)) = s(g^{-1}(\alpha)) = H\alpha$$

$$\hat{\alpha} = (H^T H)^{-1} H^T x.$$

□ مثال ۴. تخمین دامنه و فاز سینوسی

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

□ فرکانس معلوم است، دامنه و فاز را می خواهیم. در تخمین LS باید تابع زیر را کمینه کنیم:

$$J = \sum_{n=0}^{N-1} (x[n] - A \cos(2\pi f_0 n + \phi))^2$$

□ مشتق تابع فوق بر حسب A و ϕ غیر خطی است.

□ مثال ۴. ادامه

□ می دانیم: $A \cos(2\pi f_0 n + \phi) = A \cos \phi \cos 2\pi f_0 n - A \sin \phi \sin 2\pi f_0 n$

$$\alpha_1 = A \cos \phi$$

$$\alpha_2 = -A \sin \phi,$$

□ اگر قرار دهیم:

□ خواهیم داشت:

$$s[n] = \alpha_1 \cos 2\pi f_0 n + \alpha_2 \sin 2\pi f_0 n.$$

$$\mathbf{s} = \mathbf{H}\boldsymbol{\alpha}$$

$$\mathbf{H} = \begin{bmatrix} 1 & 0 \\ \cos 2\pi f_0 & \sin 2\pi f_0 \\ \vdots & \vdots \\ \cos 2\pi f_0(N-1) & \sin 2\pi f_0(N-1) \end{bmatrix},$$

تخمین LS غیر خطی

27

□ مثال ۴. ادامه

□ بنابراین: $\hat{\alpha} = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{x}$

□ و در نهایت:

$$\begin{aligned} \hat{\theta} &= \begin{bmatrix} \hat{A} \\ \hat{\phi} \end{bmatrix} \\ &= \begin{bmatrix} \sqrt{\hat{\alpha}_1^2 + \hat{\alpha}_2^2} \\ \arctan\left(\frac{-\hat{\alpha}_2}{\hat{\alpha}_1}\right) \end{bmatrix} \end{aligned}$$