



باسمہ تعالیٰ

شناسایی سیستم ها

Lecture 3

Least Square Rules

مقدمه

“Given a set of observations, which model parameters gives a model which approximates those up to the smallest sum of squared residuals?”

- قاعده کمترین مربعات (LS) خطا، یکی از مباحث پایه در شناسایی و تخمین است و این امر موجب گشته در تمامی دور ههای آموزشی مرتبط با حوزه تخمین، مورد توجه ویژه قرار گیرد.
- در این بخش تحقیقات صورت گرفته این حوزه و روش های عددی مطرح شده در آن را مرور خواهیم کرد.

روش ها تخمین مبتنی بر بهینه سازی

- یکی از ایده های که پشتوانه بسیاری از روش های تخمین (شناسایی) است، عبارت است از:

We choose model parameters $\hat{\theta}$ which explain the data as well as possible.

بر این اساس ۳ موضوع مهم لازم است مورد توجه قرار گیرد:

(Model): What is a good set of functions $\{ f_{\theta} \}$ which gives plausible models?

(Objective): In what sense do we need (formulate) the estimate be optimal?

(Optimization): How do we solve the minimization problem?

Systems Identification
by: Dr B. Moaveni

3

روش ها تخمین مبتنی بر بهینه سازی

- لذا در حالت کلی مساله تخمین (شناسایی) به یک مساله کلی به صورت زیر تبدیل می گردد؟

$$\hat{\theta} = \arg \min_{\theta} L_i (f_{\theta}(x_i, \theta), y_i)$$

که در آن:

x_i و y_i به ترتیب داده های ورودی-خروجی،

f_{θ} مدل سیستم،

θ پارامترهای مجهول مدل،

و L_i تابع مورد نظر برای بهینه سازی است.

Systems Identification
by: Dr B. Moaveni

4

روش ها تخمین مبتنی بر بهینه سازی

بر اساس تعریف ارائه شده از تخمین بهینه:

1- LS : The archetypical Least Squares (LS) estimation problem solves

$$\hat{\theta} = \arg \min_{\theta} \sum_{i=1}^n (f_{\theta}(x_i) - y_i)^2$$

2- WLS : The Weighted Least Squares (WLS) estimation problem solves

$$\hat{\theta} = \arg \min_{\theta} \sum_{i=1}^n w_i (f_{\theta}(x_i) - y_i)^2$$

Systems Identification
by: Dr B. Moaveni

5

روش ها تخمین مبتنی بر بهینه سازی

3- TA : The Tchebychev Approximation (TA) problem

$$\hat{\theta} = \arg \min_{\theta} \max_{i=1,2,\dots,n} |f_{\theta}(x_i) - y_i|$$

4- L1 : The L1 estimation problem solves

$$\hat{\theta} = \arg \min_{\theta} \sum_{i=1}^n |f_{\theta}(x_i) - y_i|$$

5- L0 : A robust approximation problem solves

$$\hat{\theta} = \arg \min_{\theta} \sum_{i=1}^n |f_{\theta}(x_i) - y_i|_0, \quad \text{where: } |z|_0 = \begin{cases} 0 & z = 0 \\ 1 & z \neq 0 \end{cases}$$

Systems Identification
by: Dr B. Moaveni

6

روش ها تخمین مبتنی بر بهینه سازی

پر واضح است که بسته به کاربرد مد نظر هر یک از موارد فوق می توانند به منظور تخمین/ شناسایی انتخاب گردند.

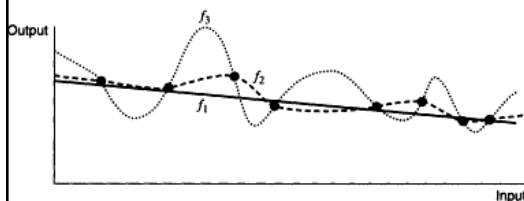
- به عنوان مثال اگر طراح به دنبال کاهش ماکزیمم قله خطا در مساله باشد، لازم است از فرموله سازی TA استفاده نماید.
- به عنوان مثال اگر طراح به دنبال آن است که تعداد نقاط بیشتری را بطور کامل بر هم منطبق نماید، لازم است از فرموله سازی L0 استفاده نماید.

Systems Identification
by: Dr B. Moaveni

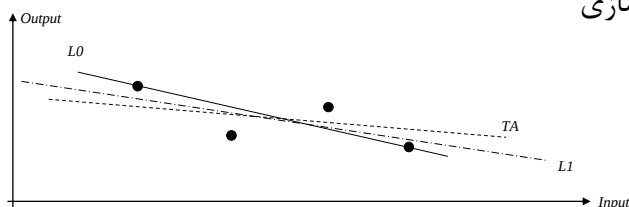
7

روش ها تخمین مبتنی بر بهینه سازی

شکل های زیر در دست یافت به اهداف مساله مورد نظر می توانند راهگشا باشند:



- انتخاب تابع یا مدل



- انتخاب روش بهینه سازی

Systems Identification
by: Dr B. Moaveni

8

روش ها تخمین مبتنی بر بهینه سازی

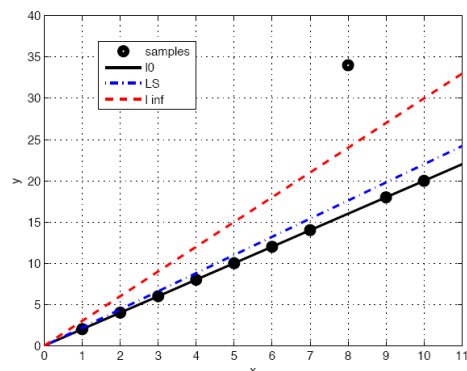


Figure 2.1: An example of an estimation problem. Assume that $\{f_\theta : \mathbb{R} \rightarrow \mathbb{R}\}$ equals the straight lines passing through the origin, with slope $\theta \in \mathbb{R}$. Assume that there are $n = 10$ samples (indicated by the black dots), with the 9th one containing a disturbance. Then the different estimators would give different solutions: (blue dashed-dotted line): the best θ according to the least squares criterion as in eq. (2.2), (red dashed line): the best θ according to the TA criterion as in eq. (2.4), (black solid line): the best θ according to the L_0 criterion as in eq. (2.6).

Systems Identification
by: Dr B. Moaveni

9

تخمین کمترین مربعات (Least Square Estimation)

روش کمترین مربعات یک راه حل کلاسیک برای تخمین (شناسایی) پارامترهای توابعی هستند که نسبت به پارامترها خطی می باشند.

• معرفی توابع خطی نسبت به پارامترها

(Linear in the Parameters (LIP))

Systems Identification
by: Dr B. Moaveni

10

معرفی توابع خطی نسبت به پارامترها

Definition 2 (LIP) A model for $\{y_i\}_i$ is linear in the unknowns $\{\theta_j\}_{j=1}^d$ if for each $y_i : i = 1, \dots, n$, one has given values $\{x_{ij}\}_{j=1}^d$ such that

$$y_i = \sum_{j=1}^d x_{ij}\theta_j + e_i, \quad (2.7)$$

and the terms $\{e_i\}$ are in some sense small. Such model can be summarized schematically as

$$\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_{11} & \dots & x_{1d} \\ \vdots & & \vdots \\ x_{n1} & \dots & x_{nd} \end{bmatrix} \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_d \end{bmatrix} + \begin{bmatrix} e_1 \\ \vdots \\ e_n \end{bmatrix}, \quad (2.8)$$

or in matrix notation

$$\mathbf{y} = \mathbf{X}\theta + \mathbf{e}. \quad (2.9)$$

where the matrix $\mathbf{X} \in \mathbb{R}^{n \times d}$, the vectors $\theta \in \mathbb{R}^d$ and $\mathbf{e} \in \mathbb{R}^n$ are as in eq. (2.8). In case d is 'small' (compared to n), one refers to them as the parameter vector.

معرفی توابع خطی نسبت به پارامترها

Example 1 (Constant Model) Perhaps the simplest example of a linear model is

$$y_t = \theta + e_t. \quad (2.10)$$

where $\theta \in \mathbb{R}$ is the single parameter to estimate. This can be written as in eq. (2.7) as

$$y_t = \mathbf{x}_t^T \theta + e_t. \quad (2.11)$$

where $\mathbf{x}_t = \mathbf{1} \in \mathbb{R}$, or in vector form as

$$\mathbf{y} = \mathbf{1}_n \theta + \mathbf{e}, \quad (2.12)$$

where $\mathbf{y} = (y_1, \dots, y_n)^T \in \mathbb{R}^n$, $\mathbf{e} = (e_1, \dots, e_n)^T \in \mathbb{R}^n$ and $\mathbf{1}_n = (1, \dots, 1)^T \in \mathbb{R}^n$. Hence, the 'inputs' take the constant value 1. This thinking is also used in order to express a model of d inputs $\mathbf{X}_t = \mathbf{x}_t^T \in \mathbb{R}^d$ for all $t = 1, \dots, n$ with a constant intercept term θ_0 given as

$$y_t = \mathbf{x}_t^T \theta + \theta_0 + e_t. \quad (2.13)$$

In matrix form one has

$$\mathbf{y} = \mathbf{X}'\theta' + \mathbf{e}, \quad (2.14)$$

where now $\theta' = (\theta^T \theta_0)^T \in \mathbb{R}^{d+1}$ and $\mathbf{X}' = [\mathbf{X} \quad \mathbf{1}_n]$.

معرفی توابع خطی نسبت به پارامترها

Example 2 (Polynomial Trend) Assume the output has a polynomial trend of order smaller than $m > 0$, then it is good practice to consider the model

$$y_t = \sum_{j=1}^d x_{tj} \theta_j + \sum_{k=0}^m t^k \theta'_k + e_t = \mathbf{x}_t^T \theta + \mathbf{z}^T(t) \theta' + e_t, \quad (2.15)$$

where $\mathbf{z}(t) = (1, t, t^2, \dots, t^m)^T \in \mathbb{R}^{m+1}$ and $\theta' = (\theta'_0, \theta'_1, \dots, \theta'_m)^T \in \mathbb{R}^{m+1}$. Again, in matrix notation one has

$$\mathbf{y} = \mathbf{X}' \theta' + \mathbf{e}. \quad (2.16)$$

where $\mathbf{X}'_t = (\mathbf{x}_t^T, 1, t, \dots, t^m)$ and $\theta' = (\theta^T, \theta'_0, \theta'_1, \dots, \theta'_m)^T \in \mathbb{R}^{d+m+1}$.

معرفی توابع خطی نسبت به پارامترها

Example 3 (A Weighted sum of Exponentials) It is crucial to understand that models which are linear in the parameters are not necessary linear in the covariates. For example, consider a nonlinear function

$$y_t = f(\mathbf{x}_t) + e_t, \quad (2.17)$$

where $f: \mathbb{R}^d \rightarrow \mathbb{R}$ is any function. Then one can find an arbitrary good approximation to this model

$$y_t = \sum_{k=1}^m \varphi_k(\mathbf{x}_t) \theta_k + e_t = \varphi(\mathbf{x}_t) \theta + e_t, \quad (2.18)$$

where $\{\phi_1, \dots, \phi_m\} \subset \{f: \mathbb{R}^d \rightarrow \mathbb{R}\}$ are an appropriate set of basis functions. Then $\varphi(\mathbf{x}) = (\varphi_1(\mathbf{x}), \dots, \varphi_m(\mathbf{x}))^T \in \mathbb{R}^m$. There are ample ways on which form of basis functions to consider. A method which often works is to work with exponential functions defined for any $k = 1, \dots, m$ as

$$\varphi_k(\mathbf{x}) = \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}_k\|}{\sigma_k}\right), \quad (2.19)$$

where $\sigma_k > 0$ and $\mathbf{x}_k \in \mathbb{R}^d$ is chosen suitably. Specific examples of basis functions are the orthogonal polynomials (e.g. Chebychev, Legendre or Laguerre polynomials). More involved sets lead to methods as wavelets, splines, orthogonal functions, or kernel methods.

معرفی توابع خطی نسبت به پارامترها

Example 5 (Nonlinear in the Parameters) Consider the following model for observations $\{y_t\}_{t=1}^n$

$$y_t = a \sin(bt + c), \quad (2.23)$$

where (a, b, c) are unknown parameters. Then it is seen that the model is linear in a , but not in b, c . A way to circumvent this is to come up with plausible values $\{b_1, \dots, b_m\}$ for b , and $\{c_1, \dots, c_m\}$ for c , and to represent the model as

$$y_t = \sum_{i,j=1}^m a_{i,j} \sin(b_i t + c_j), \quad (2.24)$$

where the model (2.23) is recovered when $a_{i,j} = a$ when $b_i = b$ and $c_j = c$, and is zero otherwise.

معرفی توابع خطی نسبت به پارامترها

Example 6 (Quadratic in the Parameters) Consider the following model for observations $\{(y_t, x_t)\}_{t=1}^n$

$$y_t = ax_t + e_t + be_{t-1}, \quad (2.25)$$

where $\{e_t\}_t$ are unobserved noise terms. Then we have cross-terms $\{be_{t-1}\}$ of unknown quantities, and the model falls not within the scope of models which are linear in the parameters.

تخمین کمترین مربعات (Least Square Estimation) (Ordinary Least Square Estimation (OLS))

تخمین کمترین مربعات یکی از روشهای قدرتمند تخمین است که بر پایه روش کمترین مربعات بنا نهاده شده است. این روش یک روش غیر تصادفی تخمین است، به عبارت دیگر در این روش اندازه گیری ها و متغیرهای مد نظر عمل تخمین نویزی نیستند.

در ادامه روش کمترین مربعات برای دو هدف زیر مد نظر قرار می گیرد:

1. تخمین پارامترهای یک تابع یا یک سیگنال.
2. تخمین متغیرهای حالت در مدل فضای حالت.

تخمین کمترین مربعات (Least Square Estimation)

یک سیگنال گسسته با توصیف زیر در نظر بگیرید:

$$(1.20) \quad \leftrightarrow \quad s(n) = \sum_{j=1}^q \theta_j \gamma_j(n) \Rightarrow s(n) = \gamma(n)\theta$$

$$\theta_1, \theta_2, \dots, \theta_q : \text{constant and unknown} \quad \theta = \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_q \end{bmatrix}$$

$\gamma_1(n), \gamma_2(n), \dots, \gamma_q(n)$: known function of n

$$\gamma(n) = [\gamma_1(n) \quad \dots \quad \gamma_q(n)]$$

Measurements of signal: $z(n) = s(n) + v(n)$

مشابه قبل هدف یافتن $\hat{\theta}(n)$ به عنوان تخمین θ است در زمان nT با استفاده از:

$$z(1), \quad z(2), \quad \dots, \quad z(n)$$

تخمین کمترین مربعات (Least Square Estimation)

با استفاده از مقدار تخمین زده شده $\hat{\theta}(n)$ میتوان سیگنال جدیدی با عنوان $\hat{s}(n)$ را ساخت که تخمینی از سیگنال واقعی $s(n)$ می باشد.

$$s(n) = \gamma(n)\theta(n), \quad (1.34)$$

همچنین مقادیر تخمینی از سیگنال را در بازه $i < n$ با استفاده از $\hat{\theta}(n)$ میتوان داشت.

$$\hat{s}(i) = \gamma(i)\hat{\theta}(n), \quad i < n \quad (1.35)$$

حال میتوان مجموع مربعاتی را به صورت زیر ارائه نمود که اگر تخمین $\hat{\theta}(n)$ با مقدار θ برابر بوده و $v(i) = 0, i = 1, 2, \dots, n$ باشد مجموع مربعات در رابطه (۱.۳۶) صفر خواهد گشت.

$$S.S. = [z(1) - \hat{s}(1)]^2 + [z(2) - \hat{s}(2)]^2 + \dots + [z(n) - \hat{s}(n)]^2 \quad (1.36)$$

Systems Identification
by: Dr B. Moaveni

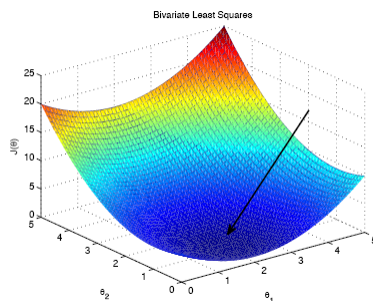
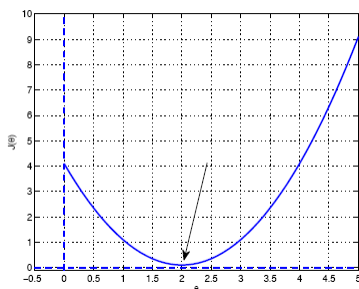
19

تخمین کمترین مربعات (Least Square Estimation)

ولیکن در صورتی که $\hat{\theta}(n)$ با θ برابر نباشد مجموع مربعات مقدار مثبتی را نتیجه خواهد داد که اندازه آن بستگی دارد به فاصله مقدار تخمین زده شده از مقدار واقعی.

لذا مجموع مربعات را میتوان به صورت یک تابع خطا در نظر گرفت که هدف حداقل نمودن مقدار آن است با استفاده از تخمین درست $\hat{\theta}(n)$.

تخمین $\hat{\theta}(n)$ که تابع مربعات خطای (۱.۳۶) را مینیمم مینماید را تخمین حداقل مربعات خطا (LS) از θ در زمان nT می نامند.



20

تخمین کمترین مربعات (Least Square Estimation)

*** به منظور محاسبه تخمین کمترین مربعات خطا (Least Square) از پارامترهای سیگنال $s(n)$ ابتدا لازم است رابطه (۱.۳۶) به صورت برداری بازنویسی گردد. به این منظور:

$$Z_n = \begin{bmatrix} z(1) \\ z(2) \\ \vdots \\ z(n) \end{bmatrix}_{n \times 1}, \quad \Gamma_n = \begin{bmatrix} \gamma(1) \\ \gamma(2) \\ \vdots \\ \gamma(n) \end{bmatrix}_{n \times q}$$

در نتیجه رابطه (۱.۳۶) را می توان به صورت زیر بازنویسی نمود:

$$S.S. = [Z_n - \Gamma_n \hat{\theta}(n)]^T [Z_n - \Gamma_n \hat{\theta}(n)]$$

تخمین کمترین مربعات (Least Square Estimation)

به منظور محاسبه تخمین $\hat{\theta}(n)$ با استفاده از حداقل سازی تابع S.S. میتوان از ایده مشتق گیری جزئی نسبت به $\hat{\theta}(n)$ استفاده نمود. در این صورت:

$$\frac{\partial(S.S.)}{\partial \hat{\theta}(n)} = -2\Gamma_n^T [Z_n - \Gamma_n \hat{\theta}(n)] = 0 \quad \rightarrow \quad \Gamma_n^T \Gamma_n \hat{\theta}(n) = \Gamma_n^T Z_n \quad (1.38)$$

شرط کافی برای محاسبه $\hat{\theta}(n)$ از این

رابطه و یا به عبارت دیگر معکوس

پذیر بودن $\Gamma_n^T \Gamma_n$ این است که

رتبه ماتریس Γ_n برابر q باشد.

$$\hat{\theta}(n) = [\Gamma_n^T \Gamma_n]^{-1} \Gamma_n^T Z_n \quad (1.39)$$

تخمین کمترین مربعات (Least Square Estimation)

چند نکته مهم:

- شرط اینکه رتبه ماتریس Γ_n برابر q باشد این است که $n \geq q$ باشد. به عبارت دیگر تعداد دسته اطلاعات اندازه گیری شده از تعداد ضرایب مجهول بیشتر باشد.
- تخمین حداقل مربعات (LS) تخمینی از θ در زمان nT ارائه می نماید که به آن روش *batch* گفته می شود، چرا که برای محاسبه آن به تمامی اطلاعات و اندازه گیری های $z(1), z(2), \dots, z(n)$ نیاز است. پر واضح است که این روش به صورت *off-line* قابل استفاده می باشد.
- شرط کامل بودن رتبه ماتریس Γ_n در اصل یک محدودیتی است بر روی تابع های $\gamma_j(n)$ به عنوان تابع های زمانی توصیف کننده سیگنال.

تخمین کمترین مربعات (Least Square Estimation)

لم: رابطه (۱.۳۸) دارای پاسخ منحصر به فردی است

$$\Gamma_n^T \Gamma_n \hat{\theta}(n) = \Gamma_n^T Z_n \quad (1.38)$$

اگر و فقط اگر بردار غیر صفری مانند $\Theta \neq 0$ وجود نداشته باشد که $\Gamma_n^T \Gamma_n \Theta = 0$.

اثبات: اگر بردار $\Theta \neq 0$ وجود داشته باشد که $\Gamma_n^T \Gamma_n \Theta = 0$ ، آنگاه به سادگی می توان نشان داد که این رابطه بی نهایت جواب خواهد داشت. به عنوان مثال، اگر $\hat{\theta}$ یک پاسخ رابطه (۱.۳۸) باشد آنگاه تمامی بردا های $\hat{\theta} + a\Theta$ (a اسکالر است) نیز پاسخ این رابطه خواهد بود.

از سوی دیگر اگر $\hat{\theta}_1$ و $\hat{\theta}_2$ دو پاسخ متفاوت رابطه (۱.۳۸) باشند آنگاه

$$\Gamma_n^T \Gamma_n (\hat{\theta}_1 - \hat{\theta}_2) = 0 \Rightarrow \Theta = \hat{\theta}_1 - \hat{\theta}_2$$

تخمین کمترین مربعات (Least Square Estimation)

مثال:

$$s(k) = 1 - 0.1k + 0.02\sqrt{k} + 1k^3 - 5k^4 = \underbrace{\begin{bmatrix} 1 & k & \sqrt{k} & k^3 & k^4 \end{bmatrix}}_{\gamma(k)} \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \\ \theta_4 \\ \theta_5 \end{bmatrix}$$

theta =

0.9934

-0.1002

0.0222

1.0000

-5.0000

با استفاده از تخمین OLS

* برای ۴۰۰ دسته داده و بدون حضور نویز

سوال: چرا متغیرهای آخر بهتر شناسایی می شوند؟

Systems Identification
by: Dr B. Moaveni

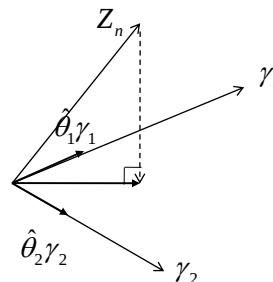
25

تخمین کمترین مربعات (Least Square Estimation)

Orthogonal Projection:

$$\Gamma_n^T [Z_n - \Gamma_n \hat{\theta}(n)] = 0 \quad \rightarrow \quad \hat{\theta}(n) = [\Gamma_n^T \Gamma_n]^{-1} \Gamma_n^T Z_n$$

$$\Gamma_n^T \left[Z_n - \underbrace{\Gamma_n \hat{\theta}(n)}_{\hat{s}} \right] = 0$$



Systems Identification
by: Dr B. Moaveni

26

تخمین کمترین مربعات (Least Square Estimation)

فرم بازگشتی روش کمترین مربعات (Recursive Least Square)

با توجه به دو علت مهم: (۱) سنگینی محاسبات در زمانی که مقدار n بزرگ می شود و (۲) نیاز به محاسبه **online** پارامترها، روش کمترین مربعات بازگشتی (RLS) ارائه گشته است.

با استفاده از معادلات زیر که درستی آنها به سادگی قابل بررسی است.

$$\Gamma_{n+1}^T Z_{n+1} = \Gamma_n^T Z_n + \gamma^T(n+1)z(n+1) \quad (1.40)$$

$$\Gamma_{n+1}^T \Gamma_{n+1} = \Gamma_n^T \Gamma_n + \gamma^T(n+1)\gamma(n+1) \quad (1.41)$$

با استفاده از قضیه معکوس سازی ماتریس:

Matrix Inversion Lemma :

$$(A_{11}^{-1} + A_{12}A_{22}A_{21})^{-1} = A_{11} - A_{11}A_{12}(A_{21}A_{11}A_{12} + A_{22}^{-1})^{-1}A_{21}A_{11}$$

$$\left[\Gamma_{n+1}^T \Gamma_{n+1} \right]^{-1} = \left[\Gamma_n^T \Gamma_n \right]^{-1} - \frac{\left[\Gamma_n^T \Gamma_n \right]^{-1} \gamma^T(n+1)\gamma(n+1) \left[\Gamma_n^T \Gamma_n \right]^{-1}}{1 + \gamma(n+1) \left[\Gamma_n^T \Gamma_n \right]^{-1} \gamma^T(n+1)} \quad (1.42) \quad 27$$

تخمین پارامترهای سیگنال با استفاده از روش RLS

با جایگزاری $n+1$ به جای n در (۱.۳۹) خواهیم داشت:

$$\hat{\theta}(n+1) = \left[\Gamma_{n+1}^T \Gamma_{n+1} \right]^{-1} \Gamma_{n+1}^T Z_{n+1} \quad (1.43)$$

با جایگزاری (۱.۴۲) و (۱.۴۰) در (۱.۴۳) خواهیم داشت:

$$\hat{\theta}(n+1) = \hat{\theta}(n) + K(n) \left[z(n+1) - \gamma(n+1)\hat{\theta}(n) \right] \quad (1.44)$$

$$K(n)_{q \times 1} = \frac{\left[\Gamma_n^T \Gamma_n \right]^{-1} \gamma^T(n+1)}{1 + \gamma(n+1) \left[\Gamma_n^T \Gamma_n \right]^{-1} \gamma^T(n+1)} \quad (1.45)$$

تخمین پارامترهای سیگنال با استفاده از روش RLS

با توجه به رابطه (۱.۴۵) و (۱.۴۲) میتوان محاسبات $K(n)$ و $[\Gamma_n^T \Gamma_n]^{-1}$ را به صورت بازگشتی انجام داد.

$$\begin{aligned} [\Gamma_{n+1}^T \Gamma_{n+1}]^{-1} &= [\Gamma_n^T \Gamma_n]^{-1} - \frac{[\Gamma_n^T \Gamma_n]^{-1} \gamma^T(n+1) \gamma(n+1) [\Gamma_n^T \Gamma_n]^{-1}}{1 + \gamma(n+1) [\Gamma_n^T \Gamma_n]^{-1} \gamma^T(n+1)} \\ &= [\Gamma_n^T \Gamma_n]^{-1} - K(n) \gamma(n+1) [\Gamma_n^T \Gamma_n]^{-1} = (I - K(n) \gamma(n+1)) [\Gamma_n^T \Gamma_n]^{-1} \end{aligned}$$

$$\begin{aligned} P_n &= [\Gamma_n^T \Gamma_n]^{-1} \\ K(n) &= \frac{P_n \gamma^T(n+1)}{1 + \gamma(n+1) P_n \gamma^T(n+1)} \\ \hat{\theta}(n+1) &= \hat{\theta}(n) + K(n) [z(n+1) - \gamma(n+1) \hat{\theta}(n)] \\ P_{n+1} &= (I - K(n) \gamma(n+1)) P_n \end{aligned}$$

Systems Identification
by: Dr B. Moaveni

29

تخمین پارامترهای سیگنال با استفاده از روش RLS

روش مربعات بازگشتی (RLS) بسیار موثرتر از روش مربعات (LS) در پروسه تخمین به کار گرفته می شود.

نکاتی مهم در خصوص استفاده از روش RLS :

- روش RLS به مقدار اولیه تخمین $\hat{\theta}$ نیاز دارد. اگر هیچ اطلاعاتی از مقدار اولیه وجود نداشته باشد می توان مقدار آن را برابر صفر در نظر گرفت.
- اولین تخمین های مناسب میتوانند در مرحله r بدست آیند. مقدار r کوچکترین مقدار صحیحی است که به ازای آن ماتریس $\Gamma_r^T \Gamma_r$ معکوس پذیر می باشد.
- مقدار ماتریس P_0 نیز لازم است در شروع الگوریتم مقدار دهی شود که پیشنهاد مناسب برای آن عبارت است از $P_0 = \alpha I, \alpha > 0$.

Systems Identification
by: Dr B. Moaveni

30

مثال

$$s(n) = s \longrightarrow \begin{cases} \gamma(n) = 1 \\ \theta = s \end{cases}$$

$$z(n) = s + v(n)$$

$$\Rightarrow \Gamma_n = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \Rightarrow \Gamma_n^T \Gamma_n = n \xrightarrow{\text{LS}} \hat{\theta}(n) = [\Gamma_n^T \Gamma_n]^{-1} \Gamma_n^T Z_n = \frac{1}{n} \sum_{i=1}^n z(i)$$

نتیجه مشابه نتیجه تخمین فیلتر میانگین است.

$$\xrightarrow{\text{RLS}} \begin{cases} K(n) = \frac{\frac{1}{n}}{1 + \frac{1}{n}} = \frac{1}{n+1} \\ \hat{\theta}(n+1) = \hat{\theta}(n) + \frac{1}{n+1} (z(n+1) - \hat{\theta}(n)) \end{cases}$$

Systems Identification
by: Dr B. Moaveni

31

روش کمترین مربعات وزن داده شده (WLS)

$$S.S. = [Z_n - \Gamma_n \hat{\theta}(n)]^T W_n [Z_n - \Gamma_n \hat{\theta}(n)] \quad \text{where, } W_n : \text{Symmetric positive definite}$$

اگر ماتریس W_n قطری باشد:

$$W_n = \begin{bmatrix} w_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & w_n \end{bmatrix}$$

$$S.S. = \sum_{i=1}^n w_i [z(i) - \hat{s}(i)]^2$$

همانگونه که مشخص است از w_i ها میتوان به عنوان ضریب فراموشی اطلاعات استفاده نمود. بطوریکه:

$$\begin{cases} w_n = 1 \\ w_i \rightarrow 0 \text{ as } i \rightarrow 1 \end{cases}$$

Systems Identification
by: Dr B. Moaveni

32

روش کمترین مربعات وزن داده شده (WLS)

Weighted Least Square:

$$\hat{\theta}(n) = [\Gamma_n^T W_n \Gamma_n]^{-1} \Gamma_n^T W_n Z_n$$

Weighted Recursive Least Square :

$$\text{If } W_n = \begin{bmatrix} \alpha^{n-1} & 0 & \dots & 0 \\ 0 & \alpha^{n-2} & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{bmatrix} \quad 0 < \alpha < 1: \text{ forgetting factor}$$

$$K(n)_{q \times 1} = \frac{[\Gamma_n^T W_n \Gamma_n]^{-1} \gamma^T(n+1)}{\alpha + \gamma(n+1) [\Gamma_n^T W_n \Gamma_n]^{-1} \gamma^T(n+1)}$$

$$[\Gamma_{n+1}^T \Gamma_{n+1}]^{-1} = (I - K(n) \gamma(n+1)) [\Gamma_n^T W_n \Gamma_n]^{-1}$$

Systems Identification
by: Dr B. Moaveni

33

روش کمترین مربعات وزن داده شده (WLS)

Recursive Least Square with Forgetting Factor:

$$\begin{aligned} K(n) &= \frac{P_n \gamma^T(n+1)}{\alpha + \gamma(n+1) P_n \gamma^T(n+1)} \\ \hat{\theta}(n+1) &= \hat{\theta}(n) + K(n) [z(n+1) - \gamma(n+1) \hat{\theta}(n)] \\ P_{n+1} &= (I - K(n) \gamma(n+1)) P_n / \alpha \end{aligned}$$

$0 < \alpha < 1$: forgetting factor

Systems Identification
by: Dr B. Moaveni

34

پروژه ۳-۱

با استفاده از روش LS و RLS پارامترهای عددی α_i تابع زیر را تخمین بزنید و مسائل زیر را بررسی کنید:

- ۱- اثر تعداد نمونه ها
- ۲- اثر مقادیر اولیه پارامترها در RLS
- ۳- اثر مقادیر اولیه ماتریس کوواریانس خطا
- ۴- اثر ضریب فراموشی
- ۵- مزایا و معایب LS و RLS نسبت به یکدیگر

$$f(t) = \alpha_0 + \alpha_1 \sin(2t) + \alpha_2 e^{-0.5t}$$

$$T_s = 0.1$$

$$\alpha_0 = 1$$

$$\alpha_1 = 0.2$$

$$\alpha_2 = -1$$

Systems Identification
by: Dr B. Moaveni

35

شناسایی یک سیستم با استفاده از روش کمترین مربعات

$$\frac{Y(z)}{U(z)} = \frac{b(z)}{a(z)} = \frac{b_{n-1}z^{n-1} + b_{n-2}z^{n-2} + \dots + b_0}{z^n + a_{n-1}z^{n-1} + \dots + a_1z + a_0}$$

$$\rightarrow \frac{Y(z^{-1})}{U(z^{-1})} = \frac{b_{n-1}z^{-1} + b_{n-2}z^{-2} + \dots + b_0z^{-n}}{1 + a_{n-1}z^{-1} + \dots + a_1z^{1-n} + a_0z^{-n}}$$

$$\rightarrow \frac{Y(z^{-1})}{U(z^{-1})} = \frac{b_{n-1}z^{-1} + b_{n-2}z^{-2} + \dots + b_0z^{-n}}{1 + a_{n-1}z^{-1} + \dots + a_1z^{1-n} + a_0z^{-n}}$$

$$\xrightarrow{I.Z.T.} y(k) = -\underset{?}{a_{n-1}}y(k-1) - \dots - \underset{?}{a_1}y(k-n+1) - \underset{?}{a_0}y(k-n) + \underset{?}{b_{n-1}}u(k-1) + \underset{?}{b_{n-2}}u(k-2) + \dots + \underset{?}{b_0}u(k-n)$$

Systems Identification
by: Dr B. Moaveni

36

شناسایی یک سیستم با استفاده از روش کمترین مربعات

$$y(k) = -\underset{?}{a_{n-1}} y(k-1) - \dots - \underset{?}{a_1} y(k-n+1) - \underset{?}{a_0} y(k-n) + \underset{?}{b_{n-1}} u(k-1) + \underset{?}{b_{n-2}} u(k-2) + \dots + \underset{?}{b_0} u(k-n)$$

$$y(k) = \begin{bmatrix} -y(k-1) & \dots & -y(k-n) & u(k-1) & \dots & u(k-n) \end{bmatrix} \begin{bmatrix} a_{n-1} \\ \vdots \\ a_0 \\ b_{n-1} \\ \vdots \\ b_0 \end{bmatrix} = \gamma(k)\theta$$

$$\theta = ?$$

Systems Identification
by: Dr B. Moaveni

37

پروژه ۲-۳

شناسایی یک سیستم حداقل مرتبه ۳

- مقایسه روش LS و RLS
- بررسی و تحلیل همگرایی پارامترها در روش RLS
 - بررسی اثر شرایط اولیه
 - بررسی مقدار اولیه ماتریس کوواریانس خطا
 - بررسی مساله بازنشانی ماتریس کوواریانس خطا
 - بررسی اثر ضریب فراموشی
- بررسی و تحلیل انتخاب مرتبه صورت و مخرج

Systems Identification
by: Dr B. Moaveni

38

تخمین حالت ها با استفاده از روش LS

$$x(n+1) = \Phi_{N \times N} x(n) \quad (1.54)$$

$$z(n) = C_{1 \times N} x(n) + v(n) \quad (1.55)$$

مشابه قبل هدف تخمین حالت ها است با استفاده از اندازه گیری های

$$z(1), z(2), \dots, z(n)$$

با فرض اینکه ماتریس Φ معکوس پذیر است معادله (۱.۵۴) را می توان در جهت عکس به صورت زیر حل نمود:

$$x(i) = \Phi^{-n+i} x(n), \quad i = 1, 2, \dots, n \quad (1.56)$$

تخمین حالت ها با استفاده از روش LS

با استفاده از روابط (۱.۵۴) تا (۱.۵۶):

$$\underbrace{\begin{bmatrix} z(1) \\ z(2) \\ \vdots \\ z(n) \end{bmatrix}}_{Z_n} = \underbrace{\begin{bmatrix} Cx(1) \\ Cx(2) \\ \vdots \\ Cx(n) \end{bmatrix}}_{V_n} + \underbrace{\begin{bmatrix} v(1) \\ v(2) \\ \vdots \\ v(n) \end{bmatrix}}_{V_n} = \underbrace{\begin{bmatrix} C\Phi^{-n+1} \\ C\Phi^{-n+2} \\ \vdots \\ C \end{bmatrix}}_{U_n} x(n) + \begin{bmatrix} v(1) \\ v(2) \\ \vdots \\ v(n) \end{bmatrix} \quad (1.57)$$

$$\rightarrow Z_n = U_n x(n) + V_n \quad (1.58)$$

در نتیجه تابع مجموع مربعات را می توان به صورت زیر نوشت:

$$S.S. = [Z_n - U_n x(n)]^T [Z_n - U_n x(n)] \quad (1.59)$$

تخمین حالت ها با استفاده از روش LS

با گرفتن مشتق جزئی از معادله (۱.۵۹) نسبت به $x(n)$ و مساوی صفر قرار دادن، تخمین متغیرهای حالت به صورت زیر قابل محاسبه است:

$$\hat{x}(n) = [U_n^T U_n]^{-1} U_n^T Z_n \quad (1.61)$$

با بازنویسی ماتریس U_n به صورت تابعی از ماتریس رؤیت پذیر O_n خواهیم داشت:

$$O_n = \begin{bmatrix} C \\ C\Phi \\ \vdots \\ C\Phi^{n-1} \end{bmatrix} : \text{observability matrix} \Rightarrow U_n = O_n \Phi^{1-n}$$

$$\rightarrow \hat{x}(n) = \Phi^{n-1} [O_n^T O_n]^{-1} O_n^T Z_n \quad (1.67)$$

تخمین حالت ها با استفاده از روش LS

توجه به معادله (۱.۶۷) نشان می دهد که در این معادله Φ^{-1} ظاهر نشده است و لذا نیازی به فرض محدود کننده معکوس پذیری Φ نیست.

$$\hat{x}(n) = \Phi^{n-1} [O_n^T O_n]^{-1} O_n^T Z_n \quad (1.67)$$

نکته قابل توجه در خصوص معادله (۱.۶۷) این است که ماتریس O_n دارای است مرتبه کامل است اگر و فقط اگر ماتریس $O_n^T O_n$ معکوس پذیر باشد و یا به عبارت دیگر مدل فضای حالت (۱.۵۴) و (۱.۵۵) رؤیت پذیر باشد.

تخمین حالت ها با استفاده از روش RLS

$$\hat{x}(n+1) = \Phi \hat{x}(n) + K(n)[z(n+1) - C\Phi \hat{x}(n)]$$

$$K(n) = \frac{P_n C^T}{1 + C P_n C^T}$$

$$P_{n+1} = \Phi [I - K(n)C] P_n \Phi^T$$

$$P_n = \Phi^n [O_n^T O_n] (\Phi^T)^n$$

خطای تخمین

خطای تخمین به صورت زیر تعریف می گردد:

$$\tilde{x}(n) = x(n) - \hat{x}(n)$$

از طرفی:

$$Z_n = U_n x(n) + V_n = O_n \Phi^{1-n} x(n) + V_n$$

در نتیجه:

$$\hat{x}(n) = \Phi^{n-1} [O_n^T O_n]^{-1} O_n^T [O_n \Phi^{1-n} x(n) + V_n]$$

$$\Rightarrow \hat{x}(n) = x(n) + \Phi^{n-1} [O_n^T O_n]^{-1} O_n^T V_n$$

$$\Rightarrow \tilde{x}(n) = x(n) - \hat{x}(n) = -\Phi^{n-1} [O_n^T O_n]^{-1} O_n^T V_n$$

این رابطه نشان می دهد که دامنه خطای تخمین ارتباط مستقیم دارد با دامنه نویز $v(n)$.