EE-602
STATISTICAL SIGNAL PROCESSING

PRESENTATION ON

CONSTRAINED LEAST SQUARES
&
NON-LINEAR LEAST SQUARES

BY:
PRATEEK TAMRAKAR
SAURABH AGRAWAL

GUIDED BY:
DR. RAJESH M HEGDE

DEPARTMENT OF ELECTRICAL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY KANPUR

CONSTRAINED LEAST SQUARES

*At times we come across the **LS** problems where unknown parameters must be constrained. Assuming parameter θ is subject to r<p independent constraints, then we summarize the constraints as*

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

*Where **A** is a **r** x **p** matrix and **b** is a known **r** x **1** vector.*
For instance, let p=2 and one parameter is negative of other, then the constraint would be

$$\theta_1 + \theta_2 = 0$$

Then we have

$$\mathbf{A} = [1 \quad 1] \quad \textit{and} \quad \mathbf{b} = 0$$

*The condition for the constraints to be independent is that the matrix **A** should be full rank (equal to **r**). There are only p-r independent parameters.*

Contd…

To find the ***L****east* ***S****quares* ***E****stimator* **(LSE)** *subject to linear constraints, we use* ***Lagrangian multiplier.***

We determine θ_c (c denotes the constrained LSE) by minimizing the Lagrangian

$$J_C=(x-H\theta)^T(x-H\theta) + \lambda^T (A\theta-b)$$

Where λ is a ***r*** x ***1*** vector of Lagrangian multipliers.

Expanding above expression, we have

$$J_C=x^Tx-2\theta^TH^Tx+\theta^TH^TH\theta+\lambda^TA\theta-\lambda^Tb$$

taking the gradient w.r.t. θ**,** we have

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

Contd…

Setting it to zero, we have

$$\hat{\theta}_C = (H^T H)^{-1} H^T x - \frac{1}{2} (H^T H)^{-1} A^T \lambda$$

$$= \hat{\theta} - (H^T H)^{-1} A^T \lambda \qquad \text{(i)}$$

Where $\hat{\theta}$ *is the unconstrained LSE.*

Now, to find λ*, we impose the constrain*

$$A\,\theta = b$$

So that

$$A\theta_C = A\,\hat{\theta} - \frac{1}{2} A (H^T H)^{-1} A^T \lambda = b$$

Hence

$$\frac{1}{2}\lambda = [A (H^T H)^{-1} A^T]^{-1} (A\theta - b)$$

Substituting in equation (i) gives

$$\hat{\theta}_C = \hat{\theta} - (H^T H)^{-1} A^T [A (H^T H)^{-1} A^T]^{-1} (A - b) \qquad \text{(ii)}$$

Where

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

Example : Constrained Signal

If the signal model is

$$\mathbf{s}[\mathbf{n}] = \begin{cases} \theta 1 \\ \theta 2 \\ 0 \end{cases}$$

and we observe **{x[0],x[1],x[2]},** *then the observation matrix is*

$$\mathbf{H} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

The signal vector is

$$\mathbf{s} = \mathbf{H}\theta = \begin{bmatrix} \theta 1 \\ \theta 2 \\ 0 \end{bmatrix}$$

Example contd…

*The unconstrained **LSE** is*

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

and the signal estimate is

$$\hat{s} = \mathbf{H}\hat{\theta} = \begin{bmatrix} x[0] \\ x[1] \\ 0 \end{bmatrix}$$

Now assume that we know a priori that

$$\theta_1 = \theta_2$$

thus we have

$$[1 \quad -1]\,\theta = 0$$

so that

$$\mathbf{A} = [1 \quad -1] \textit{ and } \mathbf{b} = 0$$

Example contd…

Now noting that

$$\mathbf{H^T H = I}$$

so from equation (ii) we have

$$\hat{\theta}_C = \hat{\theta} - \mathbf{A^T[AA^T]^{-1}A}$$
$$= [\mathbf{I} - \mathbf{A^T[AA^T]^{-1}A}]\,\hat{\theta}$$

on solving the above expression, we have

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

and the constrained signal becomes

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

Nonlinear Least Squares

Least Square Procedure estimates model parameters θ by minimizing the Least Square Error Criterion:

$$J = (x - s(\theta))^T (x - s(\theta))$$

$s(\theta)$ = Signal model for **x,** with its dependence on θ

Methods that can reduce complexity of Problem for determining nonlinear LSEs :

1. *Transformation of parameters*

2. *Separability of parameters*

Nonlinear Least Squares

1. ***Transformation of parameters***

*one-to-one transformation of θ that produces a **linear** signal model in the new space*

$$\alpha = \mathbf{g}(\theta)$$

***g** : p-dimensional function of θ whose inverse exists. If a **g** can be found so that*

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

then the signal model will be linear in α.

Nonlinear Least Squares

*Thus the **nonlinear** LSE of θ by*

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

Where : $\alpha' = (H^T H)^{-1} H^T x$

Approach: *minimization can be carried out in any transformed space that is obtained by a one-to-one mapping & then converted back to original space.*

Nonlinear Least Squares

2 . *Separability of parameters*

A separable signal model has the form

$$s = H(\alpha)\beta$$

where

$$\theta = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} (p-q)\times 1 \\ q\times 1 \end{bmatrix}$$

$H(\alpha)$ is an $N \times q$ matrix dependent on α.

This model is linear in β but nonlinear in α. The LS Error may be minimized with respect to β and thus reduced to a function of α only.

Nonlinear Least Squares

Since

$$J(\alpha,\beta) = (x - H(\alpha)\beta)^T (x - H(\alpha)\beta)$$

the β *that minimizes J for a given* α *is*

$$\beta' = (H^T(\alpha)H(\alpha))^{-1} H^T(\alpha)x$$

And the resulting LS error is

$$J(\alpha,\beta') = x^T[\, I - H(\alpha)(H^T(\alpha)H(\alpha))^{-1} H(\alpha)^T \,]x$$

The Problem now reduces to a ***maximization*** *of*

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

over α.

Nonlinear Least Squares

Example : Sinusoidal Parameter Estimation

For a sinusoidal signal model

$$s[n] = A\cos(2\pi f_o n + \phi)$$

It is desired to estimate the amplitude A, where $A > 0$, phase ϕ.

The frequency f_o is assumed known.

LSE is obtained by minimizing

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

Nonlinear Least Squares

Over A & ϕ, *a nonlinear LS Problem. However , because*

$$A\cos(2\pi f_o n + \phi) = A\cos(\phi)\cos(2\pi f_o n) - A\sin(\phi)\sin(2\pi f_o n)$$

If we let $\alpha_1 = A\cos(\phi)$

$\alpha_2 = -A\sin(\phi)$

Then the signal model becomes

$$s[n] = \alpha_1 \cos(2\pi f_o n) + \alpha_2 \sin(2\pi f_o n)$$

In the Matrix form this is $s = H\alpha$

Nonlinear Least Squares

where

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

Which is now linear in the new parameters.

The LSE of α *is*

$$\alpha' = (H^T H)^{-1} H^T x$$

and to find θ' *we must find the inverse transformation* $g^{-1}(\alpha)$.

Nonlinear Least Squares

This is

$$A = \sqrt{(\alpha_1{}^2 + \alpha_2{}^2)}$$

$$\phi = \tan^{-1}\left(\frac{-\alpha_2}{\alpha_1}\right)$$

so that the nonlinear LSE for this problem is given by

$$\theta' = \begin{bmatrix} A' \\ \phi' \end{bmatrix}$$

$$\theta' = \begin{bmatrix} \sqrt{(\alpha_1{}^2 + \alpha_2{}^2)} \\ \tan^{-1}\left(\frac{-\alpha_2}{\alpha_1}\right) \end{bmatrix}$$

where $\alpha' = (H^T H)^{-1} H^T x$