Linear Regression

G V V Sharma*

Contents

***Abstract*—This manual provides an introduction to linear lethods in regression.**

1 Steinhart–Hart Equation

1.1 The Steinhart–Hart equation is a model of the resistance of a thermistor at different temperatures. The equation is given by

$$\frac{1}{\tau} = w_1 + w_2 \ln(R) + w_3 [\ln(R)]^3 \tag{1.1}$$

Let

$$\mathbf{x}_1 = \begin{pmatrix} 1 \\ \ln(R_1) \\ [\ln(R_1)]^3 \end{pmatrix} \tag{1.2}$$

$$y_1 = \frac{1}{\tau_1} \tag{1.3}$$

$$\mathbf{w} = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} \tag{1.4}$$

Show that

$$y_1 = \mathbf{x}_1^T \mathbf{w} \tag{1.5}$$

1.2 Suppose for $n > 3$

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \mathbf{X}^T = \begin{pmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \vdots & \mathbf{x}_n \end{pmatrix}, \tag{1.6}$$

show that

$$\mathbf{y} = \mathbf{X}\mathbf{w} \tag{1.7}$$

1.3 For $\tau = 10^\circ C - 100^\circ C$, use the PT-100 resistance table in

https://github.com/gadepall/EE1390/blob/master/refs/5pt100sensoren_e.pdf?raw=true

to estimate $\mathbf{w}$ using the relation

$$\hat{\mathbf{w}} = \left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T \mathbf{y} \tag{1.8}$$

1.4 Verify your result by finding the temperature when the resistance is 175.86Ω.

2 Least Squares

2.1 Find

$$\|\mathbf{y} - \mathbf{X}\mathbf{w}\|^2 \tag{2.1}$$

Solution:

$$\|\mathbf{y} - \mathbf{X}\mathbf{w}\|^2 = (\mathbf{y} - \mathbf{X}\mathbf{w})^T (\mathbf{y} - \mathbf{X}\mathbf{w}) \tag{2.2}$$

$$= \|\mathbf{y}\|^2 - \mathbf{w}^T \mathbf{X}^T \mathbf{y} \tag{2.3}$$

$$- \mathbf{y}^T A \mathbf{w} + \mathbf{w}^T \mathbf{X}^T \mathbf{X}\mathbf{w} \tag{2.4}$$

2.2 Assuming 2×2 matrices and 2×1 vectors, show that

$$\frac{\partial}{\partial \mathbf{w}} \mathbf{w}^T \mathbf{X}^T \mathbf{y} = \frac{\partial}{\partial \mathbf{w}} \mathbf{y}^T \mathbf{X}\mathbf{w} = \mathbf{y}^T \mathbf{X} \tag{2.5}$$

2.3 Show that

$$\frac{\partial}{\partial \mathbf{w}} \mathbf{w}^T \mathbf{X}^T \mathbf{X}\mathbf{w} = 2\mathbf{w}^T \left(\mathbf{X}^T \mathbf{X}\right) \tag{2.6}$$

*The author is with the Department of Electrical Engineering, Indian Institute of Technology, Hyderabad 502285 India e-mail: gadepall@iith.ac.in.

2.4 Show that

$$\hat{\mathbf{w}} = \min_{\mathbf{w}} ||\mathbf{y} - \mathbf{X}\mathbf{w}||^2 \tag{2.7}$$

$$= \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{y} \tag{2.8}$$

2.5 Using the Gram-Schmidt orthogonalization procedure, show that

$$\mathbf{X} = \mathbf{Q}\mathbf{R} \tag{2.9}$$

where $\mathbf{Q}^T\mathbf{Q} = \mathbf{I}$ and $\mathbf{R}$ is upper triangular.

2.6 Show that

$$\hat{\mathbf{w}} = \mathbf{R}\mathbf{Q}^T\mathbf{y} \tag{2.10}$$

2.7 Find $\hat{\mathbf{y}}$

3 Ridge Regression

3.1 The ridge problem is defined as

$$\hat{\mathbf{w}} = \min_{\mathbf{w}} ||\mathbf{y} - \mathbf{X}\mathbf{w}|| \tag{3.1}$$

$$\text{s.t } ||\mathbf{w}||^2 \le t \tag{3.2}$$

Using the Lagrangian, show that

$$\hat{\mathbf{w}} = \left(\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I}\right)^{-1}\mathbf{X}^T\mathbf{y} \tag{3.3}$$

4 The Lasso

4.1 The Lasso is defined as

$$\hat{\mathbf{w}} = \min_{\mathbf{w}} ||\mathbf{y} - \mathbf{X}\mathbf{w}|| \tag{4.1}$$

$$\text{s.t } \sum_i |w_i| \le t \tag{4.2}$$

Obtain the corresponding Lagrangian.

4.2 Show that this is a quadratic programming problem and find a suitable algorithm.