

# Linear Regression

G V V Sharma\*

## CONTENTS

1	Least Squares	1
2	Ridge Regression	1
3	The Lasso	1

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

## 1 LEAST SQUARES

1.1 Find

$$\|y - Xw\|^2 \quad (1.1)$$

**Solution:**

$$\|y - Xw\|^2 = (y - Xw)^T (y - Xw) \quad (1.2)$$

$$= \|y\|^2 - w^T X^T y \quad (1.3)$$

$$- y^T A w + w^T X^T X w \quad (1.4)$$

1.2 Assuming  $2 \times 2$  matrices and  $2 \times 1$  vectors, show that

$$\frac{\partial}{\partial w} w^T X^T y = \frac{\partial}{\partial w} y^T X w = y^T X \quad (1.5)$$

1.3 Show that

$$\frac{\partial}{\partial w} w^T X^T X w = 2w^T (X^T X) \quad (1.6)$$

1.4 Show that

$$\hat{w} = \min_w \|y - Xw\|^2 \quad (1.7)$$

$$= (X^T X)^{-1} X^T y \quad (1.8)$$

1.5 Using the Gram-Schmidt orthogonalization procedure, show that

$$X = QR \quad (1.9)$$

where  $Q^T Q = I$  and  $R$  is upper triangular.

1.6 Show that

$$\hat{w} = RQ^T y \quad (1.10)$$

1.7 Find  $\hat{y}$

## 2 RIDGE REGRESSION

2.1 The ridge problem is defined as

$$\hat{w} = \min_w \|y - Xw\| \quad (2.1)$$

$$\text{s.t } \|w\|^2 \leq t \quad (2.2)$$

Using the Lagrangian, show that

$$\hat{w} = (X^T X + \lambda I)^{-1} X^T y \quad (2.3)$$

## 3 THE LASSO

3.1 The Lasso is defined as

$$\hat{w} = \min_w \|y - Xw\| \quad (3.1)$$

$$\text{s.t } \sum_i |w_i| \leq t \quad (3.2)$$

Obtain the corresponding Lagrangian.

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

\*The author is with the Department of Electrical Engineering, Indian Institute of Technology, Hyderabad 502285 India e-mail: gadepall@iith.ac.in. All content in this manual is released under GNU GPL. Free and open source.