Linear Classification

G V V Sharma*

Contents

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

1 Least Squares

1.1 Find

$$\|\mathbf{y}-\mathbf{X}\mathbf{w}\|^2 \tag{1.1}$$

Solution:

$$\|\mathbf{y}-\mathbf{X}\mathbf{w}\|^2 = (\mathbf{y}-\mathbf{X}\mathbf{w})^T(\mathbf{y}-\mathbf{X}\mathbf{w}) \tag{1.2}$$

$$= \|\mathbf{y}\|^2 - \mathbf{w}^T\mathbf{X}^T\mathbf{y} \tag{1.3}$$

$$- \mathbf{y}^T A\mathbf{w} + \mathbf{w}^T\mathbf{X}^T\mathbf{X}\mathbf{w} \tag{1.4}$$

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

$$\frac{\partial}{\partial \mathbf{w}}\mathbf{w}^T\mathbf{X}^T\mathbf{y} = \frac{\partial}{\partial \mathbf{w}}\mathbf{y}^T\mathbf{X}\mathbf{w} = \mathbf{y}^T\mathbf{X} \tag{1.5}$$

1.3 Show that

$$\frac{\partial}{\partial \mathbf{w}}\mathbf{w}^T\mathbf{X}^T\mathbf{X}\mathbf{w} = 2\mathbf{w}^T\left(\mathbf{X}^T\mathbf{X}\right) \tag{1.6}$$

1.4 Show that

$$\hat{\mathbf{w}} = \min_{\mathbf{w}}\|\mathbf{y}-\mathbf{X}\mathbf{w}\|^2 \tag{1.7}$$

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

1.5 Using the Gram-Schmidt orthogonalization procedure, show that

$$\mathbf{X} = \mathbf{Q}\mathbf{R} \tag{1.9}$$

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

1.6 Show that

$$\hat{\mathbf{w}} = \mathbf{R}\mathbf{Q}^T\mathbf{y} \tag{1.10}$$

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

2 Ridge Regression

2.1 The ridge problem is defined as

$$\hat{\mathbf{w}} = \min_{\mathbf{w}}\|\mathbf{y}-\mathbf{X}\mathbf{w}\| \tag{2.1}$$

$$\text{s.t } \|\mathbf{w}\|^2 \le t \tag{2.2}$$

Using the Lagrangian, show that

$$\hat{\mathbf{w}} = \left(\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I}\right)^{-1}\mathbf{X}^T\mathbf{y} \tag{2.3}$$

3 The Lasso

3.1 The Lasso is defined as

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

$$\text{s.t } \sum_i |w_i| \le t \tag{3.2}$$

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.