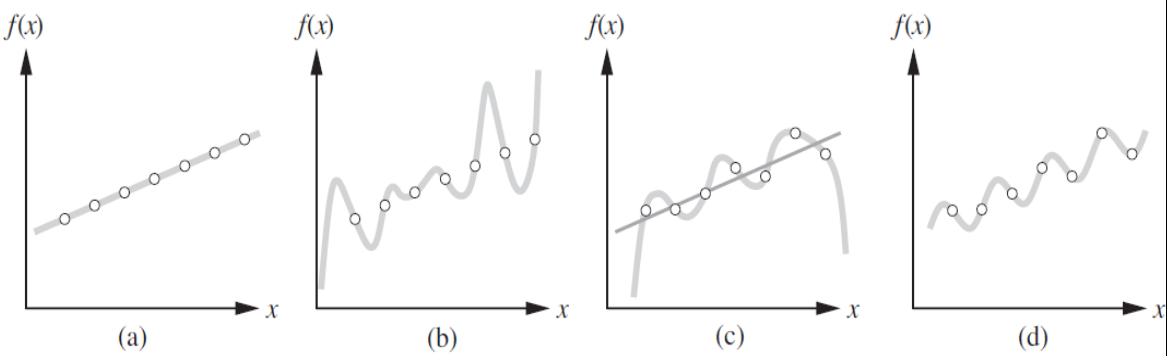
REGRESSION ANALYSIS

REGRESSION ANALYSIS

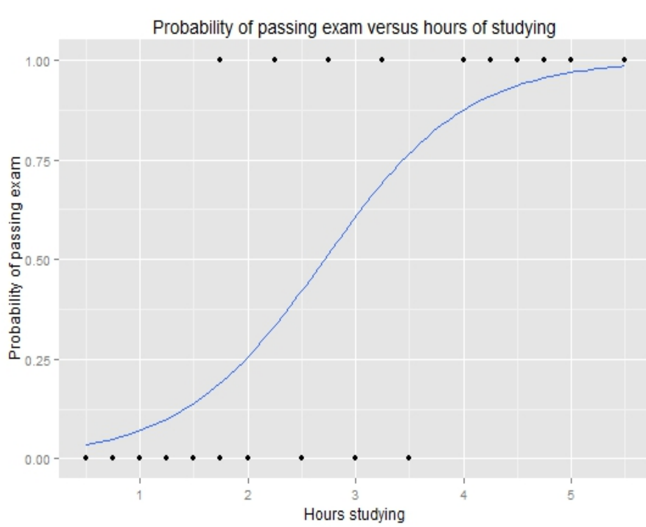
- Key idea:
  - Fitting some functions to the dataset, and trying to minimize the error.
- Definition
  - Try to find relationships between one dependent variable and independent variables
  - Think of this as f(x) = y
    - Y is dependent variable
    - Xs are independent variable

LINEAR REGRESSION



LOGISTIC REGRESSION

Retrieved from: https://upload.wikimedia.org/wikipedia/commons/6/6d/Exam_pass_logistic_curve.jpeg



LINEAR REGRESSION

LINEAR EQUATION

Linear Equation of 2D space is

- $Y = aX + b$

Linear Equation of nth-dimension spaces is

- $Y = aX_1 + bX_2 + cX_3 + \ldots + zX_n$

LINEAR REGRESSION

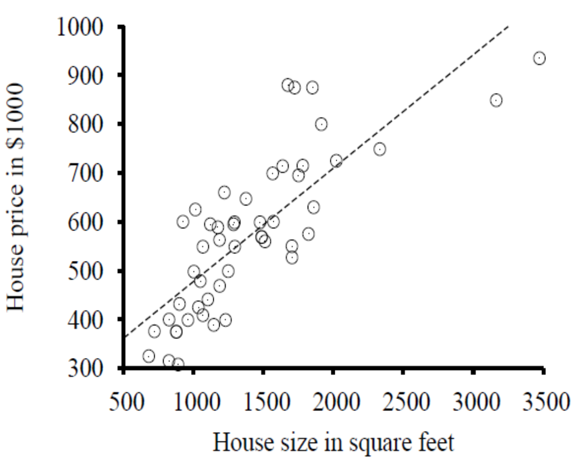
Key idea:

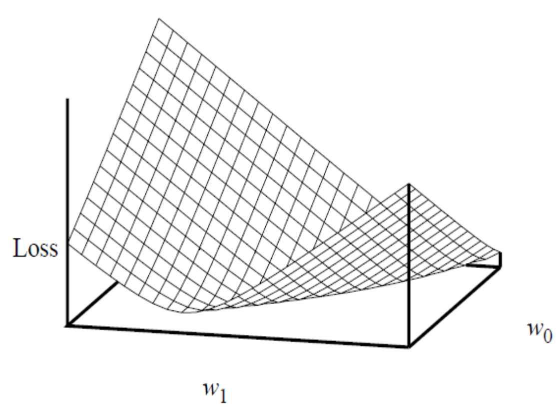
- Given features, try to find a linear equation, so it can output the expected result

Definition:

- Model of the linear relationship between a dependent variable and one or more explanatory variables

LINEAR REGRESSION





LINEAR REGRESSION

How does it work?

-Rewrite the nth-Dimension equation using Linear Algebra

$$Y(x_j) = AX_j + a_0$$

$$A = \begin{bmatrix} a_1 & a_2 & \dots & a_n \end{bmatrix}^T, \quad X_j = \begin{bmatrix} x_{j,1} & x_{j,2} & \dots & x_{j,n} \end{bmatrix}$$

LOSS FUNCTION

Define a loss function

- Sum of Squared Errors

$$Loss(Y(x_j)) = \sum_{j=1}^{N} (y_j - Y(x_j))^2$$

$$Loss(Y(x_j)) = \sum_{j=1}^{N} (y_j - (a_0 + a_1 x_{j,1} + a_2 x_{j,2} + \dots + a_n x_{j,n}))^2$$

LOSS MINIMIZATION

-Minimize the loss for 2D space

$$\frac{\partial}{\partial a_0} Loss(Y(x_j)) = 0 \quad and \quad \frac{\partial}{\partial a_1} Loss(Y(x_j)) = 0$$

-Minimize the loss for n-D space

$$\frac{\partial}{\partial a_n} Loss(Y(x_j)) = 0 \; ; \; for\ n = 0, 1, 2, ..., j$$

LOSS MINIMIZATION: GRADIENT DESCENT

- Update the weight

$$a_i \leftarrow a_i + \alpha \frac{\partial}{\partial \theta_i} Loss(Y(x_i))$$

  - α is Learning Rate

- Repeat the process until the weights converge

LOSS MINIMIZATION: LINEAR ALGEBRA

- Minimize the loss using Linear Algebra (Least Square Approximation)

$$Let\ X = [x_1\ x_2\ \dots\ x_i]^T\ \ and\ x_i = [x_{i,1}\ x_{i,2}\ \dots\ x_{i,j}]$$

$$A^* = (X^T X)^{-1} X^T y$$

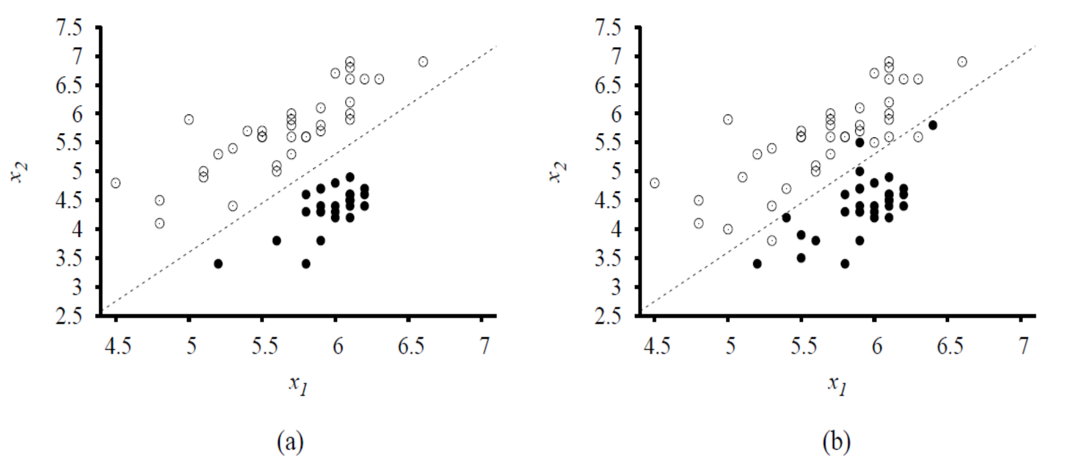


Figure 18.15 **FILES: .** (a) Plot of two seismic data parameters, body wave magnitude x_1 and surface wave magnitude x_2, for earthquakes (white circles) and nuclear explosions (black circles) occurring between 1982 and 1990 in Asia and the Middle East (?). Also shown is a decision boundary between the classes. (b) The same domain with more data points. The earthquakes and explosions are no longer linearly separable.

CLASSIFICATION PROBLEM

WHAT IS CLASSIFICATION PROBLEM?

Key idea: Classification Problem

- We have a dataset
- We try to group them into classes.

Definition:

- The process to determine if an object is a member of a set or not, or which of several sets.

CLASSIFICATION PROBLEM EXAMPLES

When you go to a supermarket

- You want to buy some chips
- Where to go?

If you want to sell tickets in a concert

- How do you price them?

CLASSIFICATION PROBLEM EXAMPLE

- *Retrieved from: http://www.imgbase.info/images/safe-wallpapers/space/star_cluster/2736-space_star_cluster_wallpaper.jpg*

LOGISTIC REGRESSION

LOGISTIC EQUATION

$$Logistic(z) = \frac{1}{1 + e^{-z}}$$

Also known as sigmoid function

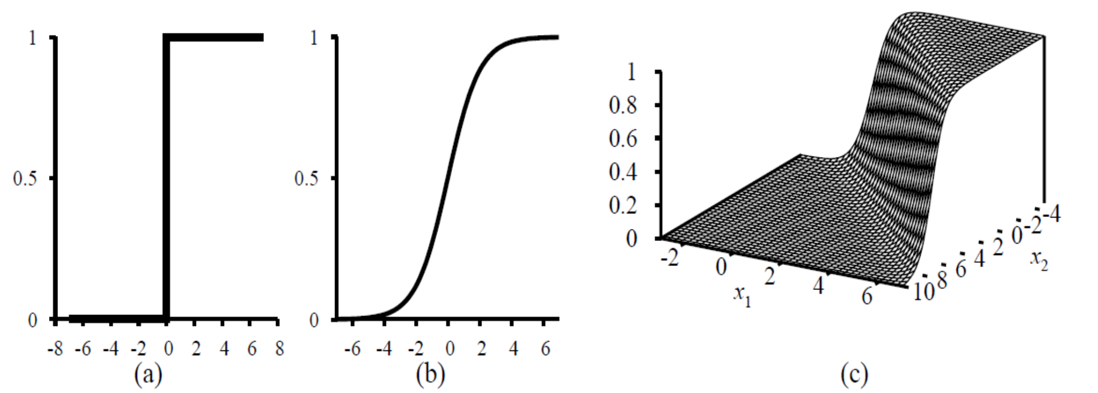


Figure 18.17 **FILES: .** (a) The hard threshold function $Threshold(z)$ with 0/1 output. Note that the function is nondifferentiable at $z = 0$. (b) The logistic function, $Logistic(z) = \frac{1}{1+e^{-z}}$, also known as the sigmoid function. (c) Plot of a logistic regression hypothesis $h_{\mathbf{w}}(\mathbf{x}) = Logistic(\mathbf{w} \cdot \mathbf{x})$ for the data shown in Figure 18.14(b).

LOGISTIC REGRESSION

Key idea:

- Regression process with logistic equation

Definition:

- The process of fitting the weights of the model (logistic model) to minimize loss on a dataset

LOGISTIC REGRESSION

- Define a loss function

$$h_w(x) = Logistic(w \cdot x) = \frac{1}{1 + e^{-w \cdot x}}$$

- Derive a loss function

$$\frac{\partial}{\partial w_i} Loss(w) = \frac{\partial}{\partial w_i} (y - h_w(x))^2$$

LOGISTIC REGRESSION

$$\frac{\partial}{\partial w_i} Loss(w) = \frac{\partial}{\partial w_i} (y - h_w(x))^2$$

$$= 2(y - h_w(x)) \times \frac{\partial}{\partial w_i} (y - h_w(x))$$

$$= 2(y - h_w(x)) \times -g'(w \cdot x) \times \frac{\partial}{\partial w_i} w \cdot x$$

$$= -2(y - h_w(x)) \times g'(w \cdot x) \times x_i$$

LOGISTIC REGRESSION

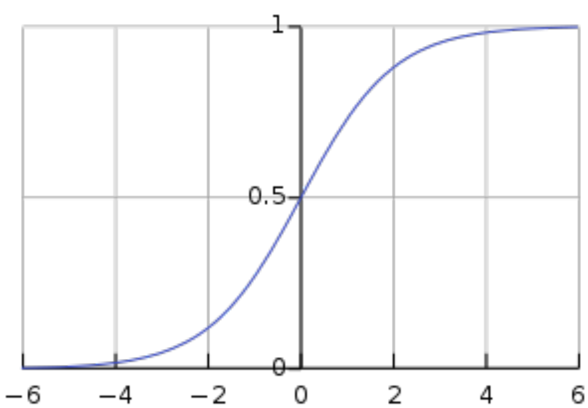
Logistic Equation has this property:

$$g'(z) = g(z)(1 - g(z))$$

Plug in, and recursively update the weight for the linear regression function.

$$w_i \leftarrow w_i + \alpha(y - h_w(x)) \times h_w(x)(1 - h_w(x)) \times x_i$$

ANALYZING LR



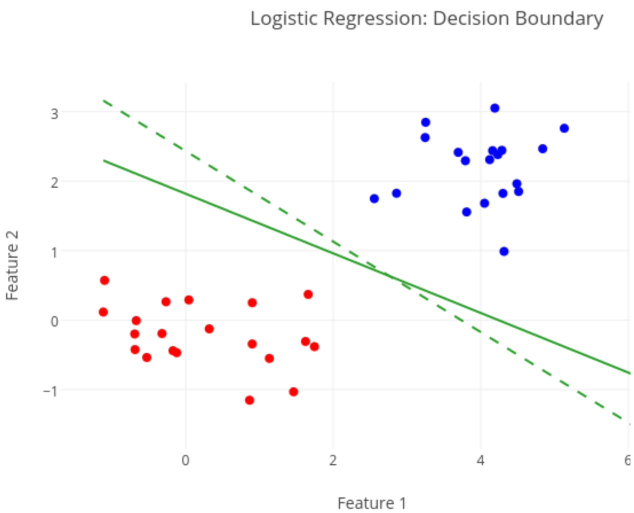
- $g(z) = 1$ if $z >= 0$
- $g(z) = 0$ if $z < 0$
- Given,
  - $z = w_0 + w_1x_1 + w_2x_2$
- It means that if

  $w_0 + w_1x_1 + w_2x_2 >= 0$

  the point (x_1,x_2) belongs to class 1

Retrieved from: https://upload.wikimedia.org/wikipedia/commons/thumb/8/88/Logistic-curve.svg/320px-Logistic-curve.svg.png

ANALYZING LR



Retrieved from: https://plot.ly/~florianh/153/logistic-regression-decision-boundary.png

- $g(z) = 1$ if $z >= 0$
- $g(z) = 0$ if $z < 0$

- Given,
  - $z = w_0 + w_1x_1 + w_2x_2$

- It means that if

$$w_0 + w_1x_1 + w_2x_2 >= 0$$

the point (x_1,x_2) belongs to class 1