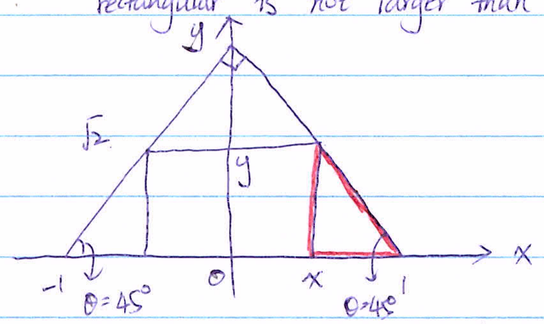
Lecture 15 (Oct. 07, 2013)

Learning Goals: (1) solve optimization problem

(2) Linear regression - Least Squares

• Geometric Optimization

Example 1: A rectangular inscribed in an isosceles right triangle whose hypotenuse is 2 units long. Find the dimensions of the rectangular that minimize its area when the length of the rectangular is not larger than 1.2 and not smaller than 0.6



Assume the length of the triangle is 2X and width is y

Area of half of the rectangular $A = x \cdot y$

Find the relation between x and y:

#1 $\sin\theta = \frac{\sqrt{2}}{2} = \frac{y}{\sqrt{(1-x)^2+y^2}}$

#2 identify the red triangle is also an isosceles & right triangle, two legs are equal: $1-x = y$

$\Rightarrow A = x(1-x) = -x^2+x$

$\Rightarrow A'(x) = -2x+1 = 0 \Rightarrow x = \frac{1}{2}$

Domain is given by $0.6 \le 2x \le 1.2 \Rightarrow 0.3 \le x \le 0.6$

$\Rightarrow \begin{cases} A(\frac{1}{2}) = \frac{1}{2} \cdot \frac{1}{2} = 0.25 \\ A(0.3) = 0.3 \times 0.7 = 0.21 \\ A(0.6) = 0.6 \times 0.4 = 0.24 \end{cases}$

$\Rightarrow x = 0.3$ is where the area is the minimum

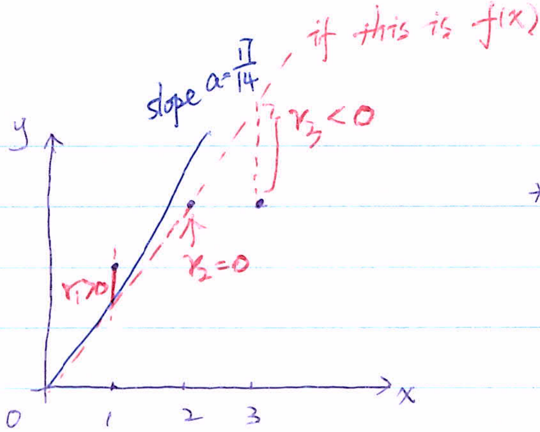
$\Rightarrow$ The dimensions of the rectangular is: 0.6 as length and 0.7 as width.

• Linear Regression - Least Square

Example 2: The hypothesis is that the photosynthesis rate is proportional to the surface area of a leaf. Suppose there's experiment which can measure the area of each leaf and their photosynthesis rates. The results are (1,2), (2,3), (3,3). x coordinate represents the surface area of a leaf, y coordinate represents the rate of photosynthesis. Can we find the straight line $y = ax$ describing the photosynthesis rate as a function of the surface area

* No such straight line can go through all three points

* Instead of saying that the hypothesis is wrong, we should notice that there are errors in the experimental data so that they can not lay on the curve perfectly



* Can we still find $y=ax$ that describes the relation of the photosynthesis rate and the surface area?

Residual: difference between the data and its expected value from the function

Suppose we have n data points $(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)$

from the function $f(x)$: $f(x_1), f(x_2), \cdots, f(x_n)$

residual of each point: $\underset{r_1}{\underset{\parallel}{y_1 - f(x_1)}}, \underset{r_2}{\underset{\parallel}{y_2 - f(x_2)}}, \cdots, \underset{r_n}{\underset{\parallel}{y_n - f(x_n)}}$

Sum of Squared Residual: $S = \sum_{i=1}^{n} r_i^2 = r_1^2 + r_2^2 + \cdots + r_n^2$

Method of Least Squares: the function that minimize S provides the best fit to the data

Back to example 2: $S = [y_1 - f(x_1)]^2 + [y_2 - f(x_2)]^2 + [y_3 - f(x_3)]^2$

$= (2-a)^2 + (3-2a)^2 + (3-3a)^2$

$= 14a^2 - 34a + 22$

$\Rightarrow S'(a) = 28a - 34 = 0 \quad \Rightarrow a = \frac{17}{14}$

$\Rightarrow S(\frac{17}{14}) = (\frac{11}{14})^2 + (\frac{8}{14})^2 + (-\frac{9}{14})^2$

If we take the domain of a as $[0, +\infty)$ we have

$S(0) = 22 > S(\frac{17}{14})$ and $\lim_{a \to \infty} S(a) = +\infty$

$\Rightarrow a = \frac{17}{14}$ provides the minimum of S