Session 7. Non-Linear Programming

* Construction of the Optimization Model

- *Decision variables*

  What does the model seek to determine?
  What are the unknown variables of the problem?

- *Objective function* (*Max* or *Min*)

  What is the objective that needs to be achieved to determine the optimum solution from among all the feasible values of the variables?

- *Constraints*

  What constraints must be imposed on the variables to satisfy the limitations of the model system?

(a) Unconstrained Optimization

- Decision variables: *Single* or *multiple* variables
- Objective function: *Single* or *multiple* objectives
  Maximize or *minimize*

(b) Constrained Optimization

- How to find the best way to allocate scarce resources?
- The resources may be raw materials, machine time or people time, money, or anything else in limited supply.
- The "best" or optimal solution may mean maximizing profits, minimizing costs, or achieving the best possible quality.

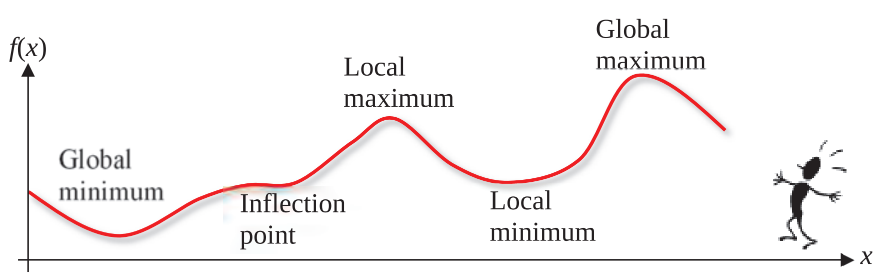
A. Univariate Optimization

* Necessary Condition

- A *necessary* condition for a particular solution $x = x_0$ to be either a *minimum* or a *maximum* is that the first-order derivative with respect to x is 0; i.e.,

$$\frac{df(x)}{dx} = 0 \text{ at } x = x_0.$$

- There are five possible solutions satisfying these conditions.



* Sufficient Condition

- To obtain more information about the five *stationary* points, it is necessary to examine the second-order derivative.

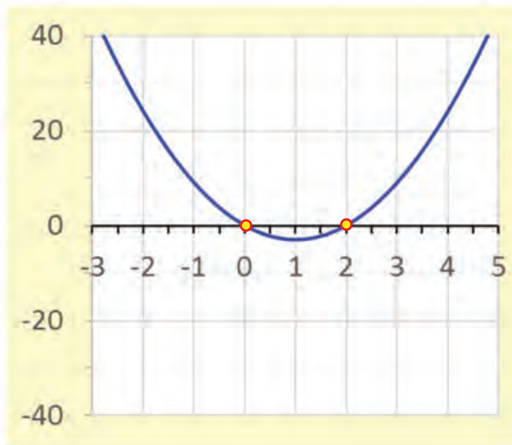
- If $\frac{d^2 f(x)}{dx^2} > 0$ at $x = x_0$, then x_0 must be *at least* a local *minimum*. If $f''(x) > 0$ for all x, then $f(x)$ is a *convex* function and x_0 is a global *minimum*.

- Similarly, if $\frac{d^2 f(x)}{dx^2} < 0$ at $x = x_0$, then x_0 must be *at least* a local *maximum*. If $f''(x) < 0$ for all x, then $f(x)$ is a *concave* function and x_0 is a global *maximum*.

Ex 1] Unconstrained Optimization

Consider the univariate function, $f(x) = x^3 - 3x^2$.

$f(x) = x^3 - 3x^2$



$f'(x) = 3x^2 - 6x$

(a) Necessary condition: The first-order derivative is

$$f'(x) = \frac{df(x)}{dx} =$$

- Thus, we have two stationary points: $x_0 =$

(b) Sufficient condition: The second-order derivative is

$$f''(x) = \frac{d^2 f(x)}{dx^2} =$$

- At $x_0 = 0$, $f''(x_0) = -6 < 0$.

  Thus, $x_0 = 0$ is a local maximum!

- At $x_0 = 2$, $f''(x_0) = 6 > 0$.

  Thus, $x_0 = 2$ is a local minimum!

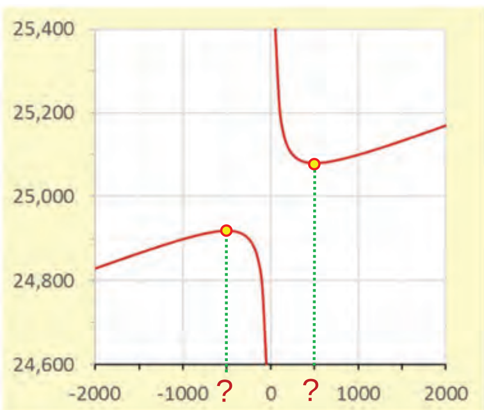
The function $f(x)$ is neither concave nor convex!

Ex 2] Minimization Problem

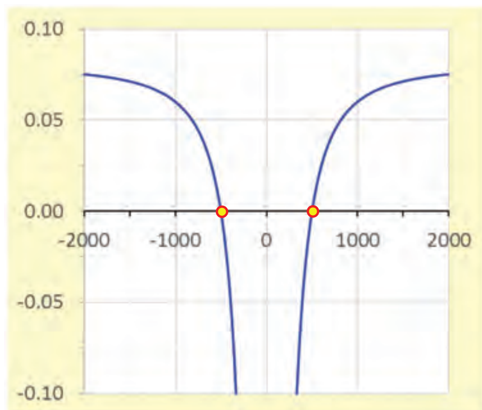
A production facility is capable of producing 60,000 widgets in a day and the total daily cost of producing x widgets in a day is given by,

$$c(x) = 25{,}000 + 0.08x + 20{,}000x^{-1}.$$

How many widgets per day should they produce in order to minimize the production cost?



$c(x)=25{,}000+0.08x+20{,}000x^{-1}$



$c'(x)= 0.08 - 20{,}000\, x^{-2}$

(a) *First-order* derivative of $c(x)$:

$$c'(x) = \frac{d}{dx}c(x) \;=$$

Thus, the stationary points are $x^* =$

(b) *Second-order* derivative of $c(x)$:

$$c''(x) = \frac{d}{dx}c'(x) \;= \quad =$$

At $x^*=+500$, $c''(x) > 0$. Thus, $x^*=+500$ is the

At $x^*=-500$, $c''(x) < 0$. Thus, $x^*=-500$ is the

(Of course, the production quantity should be $0 \le x \le 60{,}000$.)

* Inflection Point

- If the *second*-order derivative $f''(x_0)$ vanishes, i.e.,

$$\frac{d^2 f(x)}{dx^2} = 0 \quad \text{at } x = x_0,$$

then, higher-order derivatives must be investigated.

- That is, if at a stationary point x_0 of $f(x)$, the first (n-1) derivatives vanish and $f^{(n)}(x) \neq 0$, then at $x = x_0$, $f(x)$ has

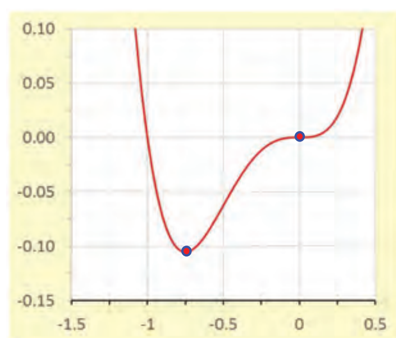
  (i) an inflection point if n is odd.

  (ii) an extreme point if n is even; i.e.,

  x_0 is a *maximum* if $f^{(n)}) < 0$

  x_0 is a *minimum* if $f^{(n)}) > 0$

Ex] Consider the univariate function, $f(x) = x^4 + x^3$.



- The first-order derivative:

  $f^{(1)}(x) = 4x^3 + 3x^2 \ = x^2(4x+3)$

  Thus, the stationary points are

  $x_0 =$

- The second-order derivative: $f^{(2)}(x) =$

  At $x_0 = -3/4$, $f^{(2)}(x_0) = 9/4 > 0$. $x_0=-3/4$ is the minimum!

  At $x_0 = 0$, $f^{(2)}(x_0) = 0$. $x_0=0$ is still unknown!

- The third-order derivative: $f^{(3)}(x) =$

  At $x_0 = 0$, $f^{(3)}(x_0) = 6 > 0$. n=3 is an *odd* number.

  Thus, $x_0=0$ is an inflection point!

B. Multivariate Optimization

* Multivariate Unconstrained Optimization

A monopolist producing a single product has two types of customer. If x_1 units are produced for customer 1, then customer 1 is willing to pay a price of (70 - 4x_1) dollars. If x_2 units are produced for customer 2, then customer 2 is willing to pay a price of (150 - 15x_2) dollars. For x_i>0, the fixed cost of manufacturing x_i units is \$100 and the variable cost is \$15. To *maximize* the profit, how much should the monopolist sell to each customer?

- Decision variables

  x_1 = units of the product sold to customer 1

  x_2 = units of the product sold to customer 2

- Objective function

  $Max\ f(x_1, x_2) =$

- Optimal solution

  $x_1^* =$ and $x_2^* =$ Then, $z^* =$

	A	B
1	x_1 =	1
2	x_2 =	1
3	$f(x_1,x_2)$ =	=B1 * (70-4*B1) + B2 * (150-15*B2) – (100 + 15 * (B1+B2))

- Objective function: Cell **B3**

- Variables: Cells **B1**:**B2**

* Multivariate Optimization with Constraints

A company is planning to spend $10,000 on advertising. It costs $3,000 per minute to advertise on television and $1,000 per minute to advertise on radio. If the firm buys x minutes of television advertising and y minutes of radio advertising, its revenue (in thousands of dollars) is given by

$$f(x, y) = xy - 2x^2 - y^2 + 8x + 3y.$$

How can the firm *maximize* its revenue?

- Objective function

  $$Max\ f(x, y) = xy - 2x^2 - y^2 + 8x + 3y$$

- Constraints

  $$3{,}000x + 1{,}000y \le 10{,}000$$

  $$x \text{ and } y \ge 0.$$

- Optimal solution

  $x^* =$ and $y^* =$ Then, $z^* =$

	A	B
1	x =	1
2	y =	1
3	f(x, y) =	=B1*B2 - 2*B1^2 - B2^2 + 8 * B1 + 3*B2
4	Constraints	=3000*B1 + 1000*B2

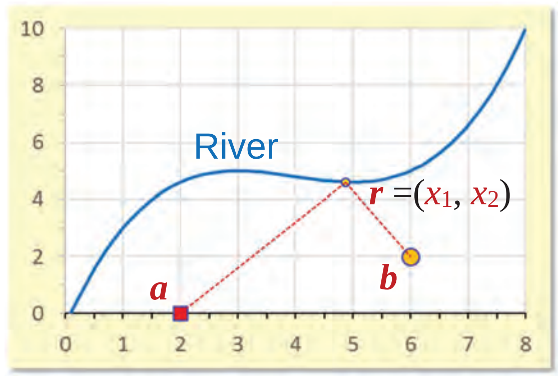
- Objective function: Cell **B3**

- Variables: Cells **B1**:**B2**

- Constraint: Cell **B4**

* Non-Linear Programming (NonLP)

A river, in the shape of a smooth curve $[(x_1-3)^3 - 3(x_1-3)^2]/10 - x_2 + 5 = 0$, flows near a house $\boldsymbol{a}$ =(2, 0) and a barn $\boldsymbol{b}$=(6, 2).



Each morning, a milkmaid leaves the house $\boldsymbol{a}$, fills a bucket of water at a point $\boldsymbol{r}$ on the river, then goes to the barn $\boldsymbol{b}$ for her cow.

Find the point $\boldsymbol{r} = (x_1, x_2)$ that *minimizes* the total distance.

- Variables

(x_1, x_2) = Coordinates of the location r

- Objective function

$$Min\ z = \sqrt{(x_1-2)^2+(x_2-0)^2} + \sqrt{(x_1-6)^2+(x_2-2)^2}$$

- Constraints

$[(x_1-3)^3 - 3(x_1-3)^2]/10 - x_2 + 5 = 0$

x_1 and x_2 are *unrestricted* in sign

- Optimal solution with Microsoft Excel - Solver

$x_1^* = 0.4170$ and $x_2^* = 1.2751$

and the minimum distance is $z^* = 7.6625$.

The milkmaid problem is a classic example of non-linear programming problem that can be solved with a *Lagrange multiplier*.

* Maximum Likelihood Estimator (*MLE*)

Consider the *exponential* random variables, x_i:

$$f(x_i) = \lambda e^{-\lambda x_i}, \qquad \text{where } i=1, 2, \ldots, n.$$

For given n sample observations, we can find the *MLE* of λ that maximizes the likelihood function $L(\lambda)$.

- Likelihood function, $L(\lambda)$

$$P[X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n]$$
$$= P[X_1 = x_1]P[X_2 = x_2]\ldots P[X_n = x_n]$$
$$= \lambda e^{-\lambda x_1} \cdot \lambda e^{-\lambda x_2} \cdot \ldots \cdot \lambda e^{-\lambda x_n} =$$

- Log-likelihood function, $ln\ L(\lambda)$

$$Max\ ln\ L(\lambda) =$$

- First-order derivative of $ln\ L(\lambda)$

$$\frac{d}{d\lambda} ln\ L(\lambda) =$$

Thus, the *MLE* of λ is $\hat{\lambda} = \frac{n}{\sum_{i=1}^{n} x_i} =$

- Second-order derivative of $ln\ L(\lambda)$

$$\frac{d^2}{d\lambda^2} ln\ L(\lambda) =$$

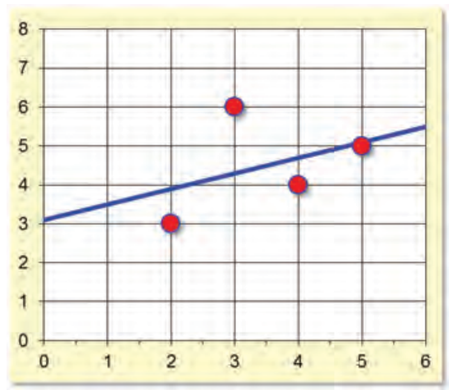
Thus, $L(\lambda)$ is a concave function and $\hat{\lambda}$ is the global maximum.

C. Prediction and Classification Models*

* Linear Regression Analysis

Find the regression parameters b_0 and b_1 that *minimize* the following performance measure:

y_i	x_i
3	2
4	4
6	3
5	5



(a) Sum of squared errors (*SSE*):

$$Min\ z = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \qquad \text{where } \hat{y}_i = b_0 + b_1 x_i$$

	A	B	C	D	E
1	Y_i	X_i	y-hat	Error	Error^2
2	3	2	3.90	-0.90	0.81
3	4	4	4.70	-0.70	0.49
4	6	3	4.30	1.70	2.89
5	5	5	5.10	-0.10	0.01
6					
7	b_0	b_1			
8	3.1	0.4		SSE =	4.20

Cell **C2** =A$8 + B2*B$8
Cell **D2** =A2-C2
Cell **E2** =D2^2
Cell **E8** =SUM(E2:E5)

- Objective function => Cell **E8**
- Decision variables => Cells **A8**:**B8**

▪ Optimal solution: b_0=3.1 and b_1=0.4 with $SSE=z^*$= 4.20

▪ Matrix representation: $\underset{4\times1}{\mathbf{Y}} = \begin{bmatrix} 3 \\ 4 \\ 6 \\ 5 \end{bmatrix}$ and $\underset{4\times2}{\mathbf{X}} = \begin{bmatrix} 1 & 2 \\ 1 & 4 \\ 1 & 3 \\ 1 & 5 \end{bmatrix}$

▪ Least Square Estimate: $\mathbf{b} = (\underset{2\times4}{\mathbf{X}'}\underset{4\times2}{\mathbf{X}})^{-1}(\underset{2\times4}{\mathbf{X}'}\underset{4\times1}{\mathbf{Y}}) =$

(b) Mean absolute deviation (*MAD*):

$$Min\ z = \frac{\sum_{i=1}^{n} |y_i - \hat{y}_i|}{n} \qquad \text{where } \hat{y}_i = b_0 + b_1 x_i$$

	A	B	C	D	E
1	Y_i	X_i	y-hat	Error	|Error|
2	3	2	3.00	0.00	0.00
3	4	4	4.32	-0.32	0.32
4	6	3	3.66	2.34	2.34
5	5	5	4.98	0.02	0.02
6					
7	b0	b1			
8	1.679	0.661		MAD =	**0.670**

Cell **C2** =A$8 + B2*B$8

Cell **D2** =A2-C2

Cell **E2** =ABS (D2)

Cell **E8** =SUM(E2:E5) / 4

- Objective function => Cell **E8**

- Decision variables => Cells **A8**:**B8**

▪ Optimal solution:

$b_0 = 1.679$ and $b_1 = 0.661$ with $MAD = z^* = 0.670$

* Logistic Regression Analysis

- Maximum likelihood estimator (for *logit* model)

$$Max\ z = \sum_{i=1}^{n} y_i \ln \pi_i + \sum_{i=1}^{n} (1 - y_i) \ln(1 - \pi_i),$$

$$\text{where } \pi_i = \frac{1}{1+exp(-\boldsymbol{\beta}'\boldsymbol{x}_i)} \quad \text{and} \quad \boldsymbol{\beta}'\boldsymbol{x}_i = b_0 + b_1 x_i$$

Ex] Find the *maximum likelihood estimates* b_0 and b_1

	A	B	C	D	E	F
1	y_i	x_i	$b_0+b_1x_i$	π_i	$y_i\ ln(\pi_i)$	$(1-y_i)\ ln(1-\pi_i)$
2	0	1	-2.556	0.072	0	-0.075
3	0	3	-1.405	0.197	0	-0.219
4	1	4	-0.829	0.304	-1.191	0
5	0	5	-0.253	0.437	0	-0.574
6	1	6	0.322	0.580	-0.545	0
7	0	7	0.898	0.711	0	-1.240
8	1	8	1.474	0.814	-0.206	0
9	1	9	2.050	0.886	-0.121	0
10	b_0	b_1			Max z =	-4.172
11	-3.132	0.576				

Cell **C2** =A$11 + B$11*B2

Cell **D2** =EXP(C2) / (1+EXP(C2))

Cell **E2** =A2 * LN(D2)

Cell **F2** =(1-A2) * LN(1-D2)

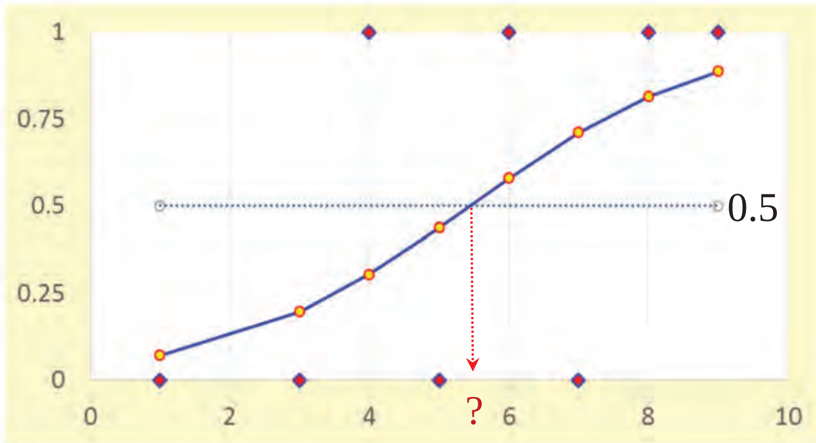
Cell **F19** =SUM(E2:F9)

- Objective function => Cell **F10**

- Decision variables => Cells **A11**:**B11**

- Solution: b_0 = -3.132 and b_1 = 0.576 with z^* = -4.171

- Logistic curve:



- Classification rule with the cutoff point $\pi^* = 0.5$

  If $X_i < -b_0/b_1 =$, assign X_i to $Y = \mathbf{0}$.

  If $X_i > -b_0/b_1 =$, assign X_i to $Y = \mathbf{1}$.

- Confusion matrix with $\pi^* = 0.5$

	Classified as 0	Classified as 1	Total
Y=0 (Not respond)	3	1	4
Y=1 (Respond)	1	3	4
Total	4	4	8

- *Error rate =*

- Suppose that the revenue is \$5 for "1" and the mailing cost is \$1:

  Expected profit =

- Best classification rule? Set the cutoff point π^* so that
  - the error rate is minimized, or
  - the expected profit is maximized.

Appendix: SAS/OR for Non-Linear Optimization

* OPTMODEL

The OPTMODEL procedure, a general purpose optimization modeling language, can also be used for concisely modeling nonlinear programming problems. Within OPTMODEL you can declare a nonlinear optimization model, pass it directly to various solvers, and review the solver result.

Ex 1] Unconstrained optimization:

Min $f(x, y) = x^2 - x - 2y - xy + y^2$

```
/* invoke procedure */
proc optmodel;
   var x, y;  /* declare variables */
   /* objective function */
   min z=x**2 - x - 2*y - x*y + y**2;
   /* now run the solver */
   solve;
   print x y;
quit;
```

Ex 2] Rosenbrock problem:

Min $f(x_1, x_2)=\alpha\,(x_2-x_1^2)^2 + (1-x_1)^2$

```
proc optmodel;
   number alpha = 100; /* declare parameter */
   var x {1..2};       /* declare variables */
   /* objective function */
   min f = alpha*(x[2] - x[1]**2)**2 +
           (1 - x[1])**2;
   /* now run the solver */
   solve;
   print x;
quit;
```