Compounding
複利

J.-S. Roger Jang (張智星)
MIR Lab, CSIE Dept.
National Taiwan University
jang@mirlab.org, http://mirlab.org/jang

2023/9/20

複利的故事

- 故事
  - 國王和棋盤
  - 一莫耳的新台幣有多少？
  - 為什麼領18%是老賊？
- 信用卡循環利率
  - George & Mary現金卡廣告（救急免利卡）
  - 2006年，台灣有70萬人淪為卡奴，平均欠款數100萬新台幣… (wiki)
  - 2010年：最新信用卡循環利率 最低2.74%最高20% (link)
  - 如果你有卡債100萬、18%計息
    - 月付2萬，約4年可還清 ➔ 錯！需要八年！
    - 月付1萬，約8年可還清 ➔ 錯！一輩子也還不完！

複利基本公式

- Basic formula

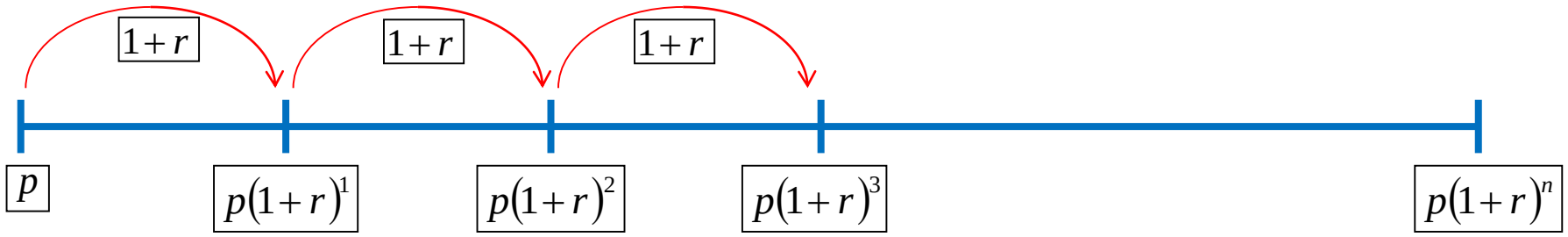
Quiz!

$f = p(1+r)^n$
p = present value (principal)
f = final value
r = interest rate per time period
n = no. of time periods

- Example
  - Initial investment: 1000
  - Annual interest: 3%
  - No. of periods: 20 years

$f = 1000 * (1 + 3/100)^{20}$
$= 1806.11$

This is NOT the formula used by banks!



定期複利（Periodic Compounding）

- 公式

$$f = p\left(1 + \frac{r}{n}\right)^{tn}$$, r: 年利率，n: 每年期數

- 範例
  - 本金1000，利率3%，計算20年後的本利和
    - 年複利

$$f = 1000 * (1 + 3\%)^{20} = 1806.11$$

    - 月複利

$$f = 1000 * \left(1 + \frac{3\%}{12}\right)^{20*12} = 1820.75$$

    - 日複利

$$f = 1000 * \left(1 + \frac{3\%}{365}\right)^{20*365} = 1822.07$$

- 提醒：
  - 我們一般所講的利率指的是年利率（r），但是
    - 銀行計算房貸是以月利率（r/12）為主
    - 信用卡循環利率計算是以日利率（r/365）為主

連續複利（Continuous Compounding）

- 公式

$$f = p * \lim_{n\to\infty}\left(1+\frac{r}{n}\right)^{tn} = p * \left(\lim_{n/r\to\infty}\left(1+\frac{1}{n/r}\right)^{n/r}\right)^{rt} = pe^{rt}$$

Euler's number, see next page.

- 範例

$$f = 1000 * e^{20*3\%} = 1822.12$$

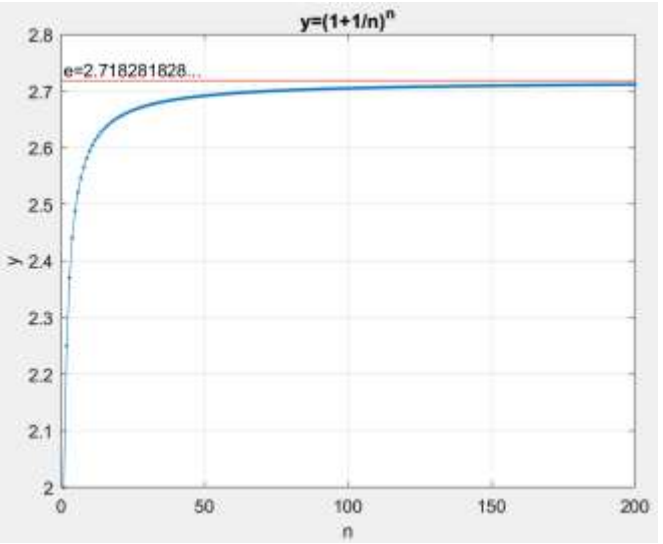
Euler's Number

- Definition

Quiz!

$$e = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = \lim_{n\to\infty}\sum_{k=0}^{n}\frac{1}{k!} = \frac{1}{0!}+\frac{1}{1!}+\frac{1}{2!}+\frac{1}{3!}+\cdots$$

=2.718281828…



- Proof by binomial theorem (二項式定理)

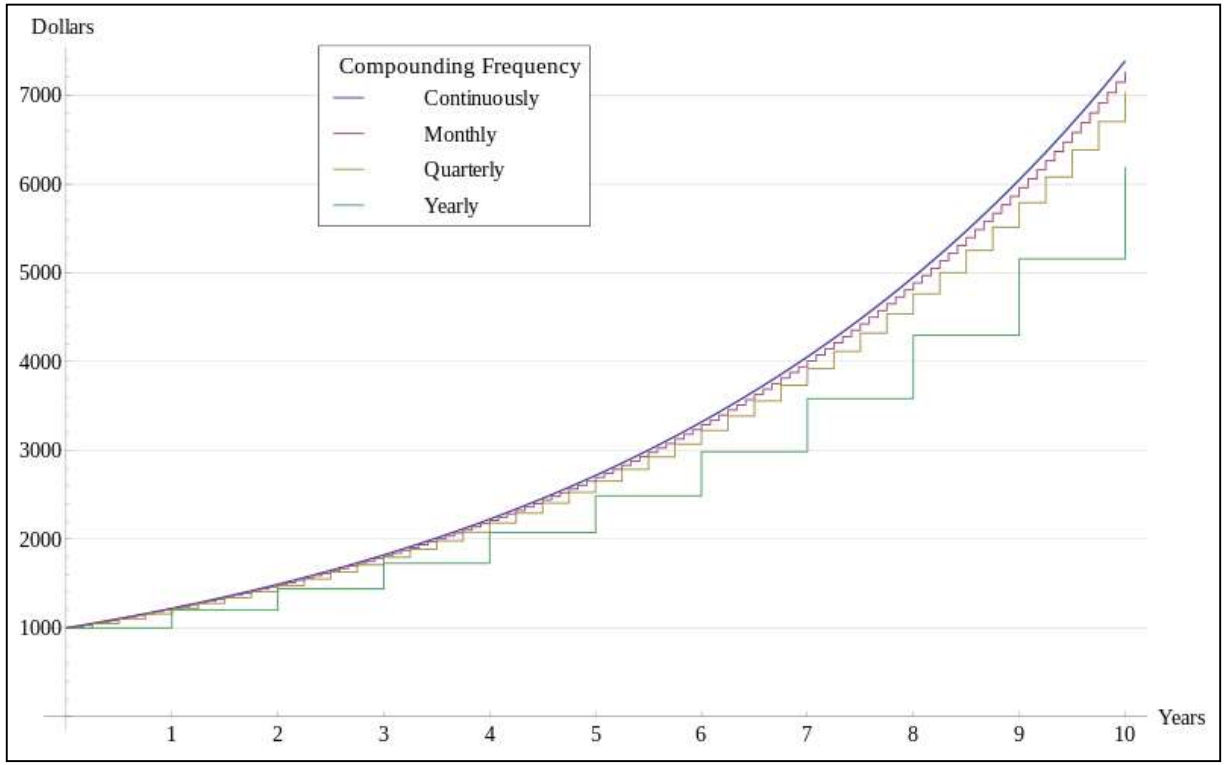
Quiz!

$$(x+y)^n = \sum_{k=0}^{n} C_k^n x^{n-k}y^k = \sum_{k=0}^{n}\frac{n!}{(n-k)!\,k!}x^{n-k}y^k$$

$$= x^n + nx^{n-1}y + \frac{n(n-1)}{2}x^{n-2}y^2 + \frac{n(n-1)(n-2)}{6}x^{n-3}y^3 + \cdots + y^n$$

Comparison

- Effect of compounding at various frequencies, with an initial investment of 1000 and 20% annual interest. (wiki)



Rule of 70

- Goal
  - To estimate the number of years it would take for an investment to double
- Also known as
  - Rule of 72
  - Rule of 69
- Reference
  - Wiki
- Formula
  - T*r=70

Quiz!

    - T=no. of year to double an investment
    - r=annual interest rate (%)
- Example:
  - r=1% ➔ T=70
  - r=3% ➔ T=23.3
  - r=18% ➔ T=3.9
  - r=20% ➔ T=3.5

Proof of Rule of 70

- 3 Ways of compounding to derive the rule

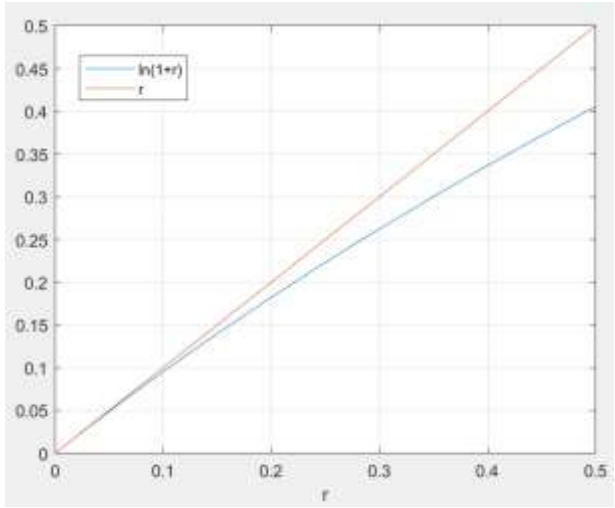
Yearly compounding: $f/p = 2 = (1+r)^T \Rightarrow T_1 = \dfrac{\ln(2)}{\ln(1+r)} \approx \dfrac{\ln(2)}{r} = \dfrac{0.6931}{r} \approx \dfrac{0.7}{r}$

Monthly compounding: $f/p = 2 = \left(1+\dfrac{r}{12}\right)^{12T} \Rightarrow T_2 = \dfrac{\ln(2)/12}{\ln(1+r/12)} \approx \dfrac{\ln(2)/12}{r/12} = \dfrac{0.6931}{r} \approx \dfrac{0.7}{r}$

Continuous compounding: $f/p = 2 = e^{rT} \Rightarrow T_3 = \dfrac{\ln(2)}{r} = \dfrac{0.6931}{r} \approx \dfrac{0.7}{r}$

- Examples

$$\mathrm{r} = 3\% \Rightarrow \begin{cases} \mathrm{T}_1 = 23.45 \approx 70/3 = 23.33 \\ \mathrm{T}_2 = 23.13 \approx 70/3 = 23.33 \\ \mathrm{T}_3 = 23.10 \approx 70/3 = 23.33 \end{cases}$$

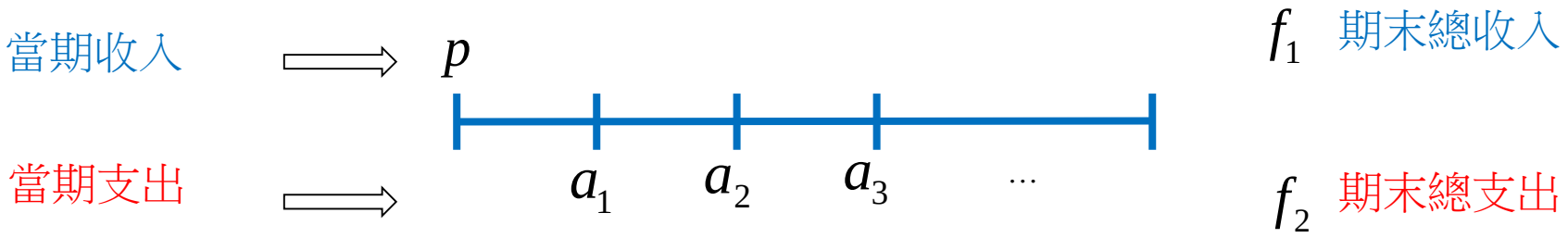


Quiz!

Prove that $\lim_{r\to 0} \dfrac{\ln(1+r)}{r} = 1.$

貨幣時間價值

- 貨幣時間價值 (Time value of money)
  - 若投資得當，金錢的價值通常會隨著時間的推移而增加。(The value of money usually increases over time if invested properly.)
- 現金流量表 (Cash flow statement)
  - 在時間軸上秀出收入及支出，並以期末（或期初）總值來判斷投資是否划算



房貸計算

- 問題定義
  - 銀行貸款100萬，20年還清，利率固定為3%，請問每個月還款金額？
- 銀行傳統
  - 利率 ➔ 年利率
  - 複利計算方式 ➔ 以「月」為單位來計算複利
- 銀行政策
  - 因為有不動產擔保品（房子），所以利率特別低
    - 若屋主無法按時繳款，銀行可以收回房子並拍賣
  - 信用貸款則屬於無擔保品的貸款，風險較高，所以利率也高
    - 若貸款方無法按時繳款，銀行可以扣此人的薪水

房貸攤還的方式

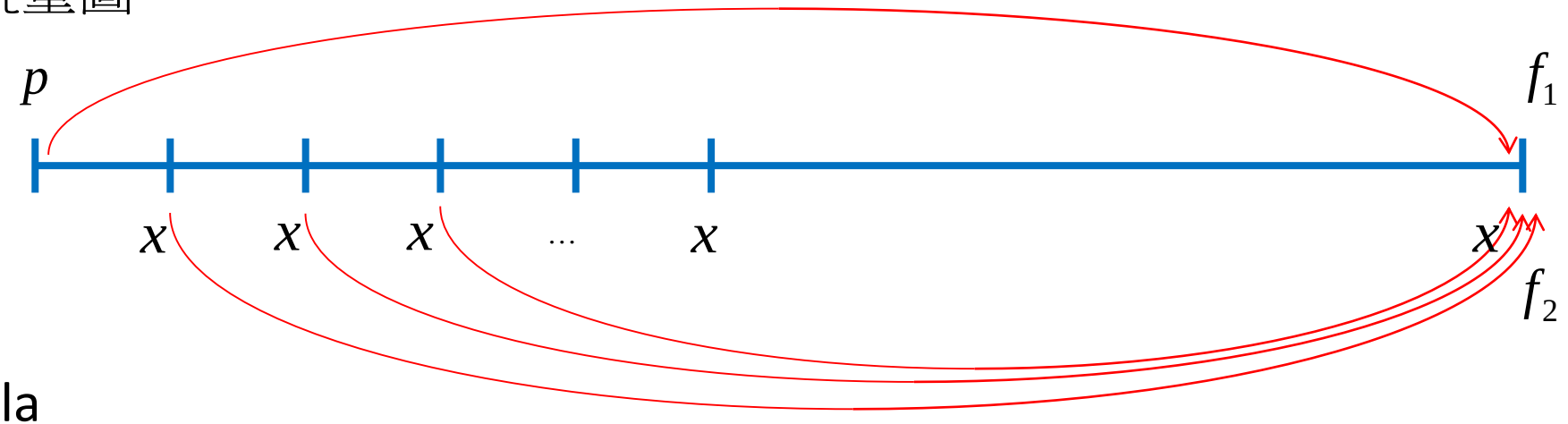
- 兩種房貸攤還的方式
  - 本息平均攤還
  - 本金平均攤還
- 相關資訊
  - 房貸試算器
  - 買房前，這些貸款名詞要知道！
  - 房貸怎麼還最有利？等額本金VS等額本息選哪個好？

等額本息與等額本金還款比較表（以貸款1萬元1年月利率為3.45‰為例）

貸款年限	等額本息			等額本金		
	還款本金	利息	本息合計	還款本金	利息	本息合計
1	817.64	34.50	852.14	833.33	34.50	867.83
2	820.46	31.68	852.14	833.33	31.63	864.96
3	823.29	28.85	852.14	833.33	28.75	862.08
4	826.13	26.01	852.14	833.33	25.88	859.21
5	828.98	23.16	852.14	833.33	23.00	856.33
6	831.84	20.30	852.14	833.33	20.13	853.46
7	834.71	17.43	852.14	833.33	17.25	850.58
8	837.59	14.55	852.14	833.33	14.38	847.71
9	840.48	11.66	852.14	833.33	11.50	844.83
10	843.38	8.76	852.14	833.33	8.63	841.96
11	846.29	5.85	852.14	833.33	5.75	839.08
12	849.21	2.93	852.14	833.37	2.88	836.25
合計	10000.00	225.68	10225.68	10000.00	224.28	10224.28

本息平均攤還：如何計算月付額

- 現金流量圖



- Formula

$$\begin{cases} f_1 = p(1+r/12)^{12*20} \\ f_2 = x + x(1+r/12) + x(1+r/12)^2 + \cdots + x(1+r/12)^{239} = x\dfrac{(1+r/12)^{240}-1}{r} \end{cases}$$

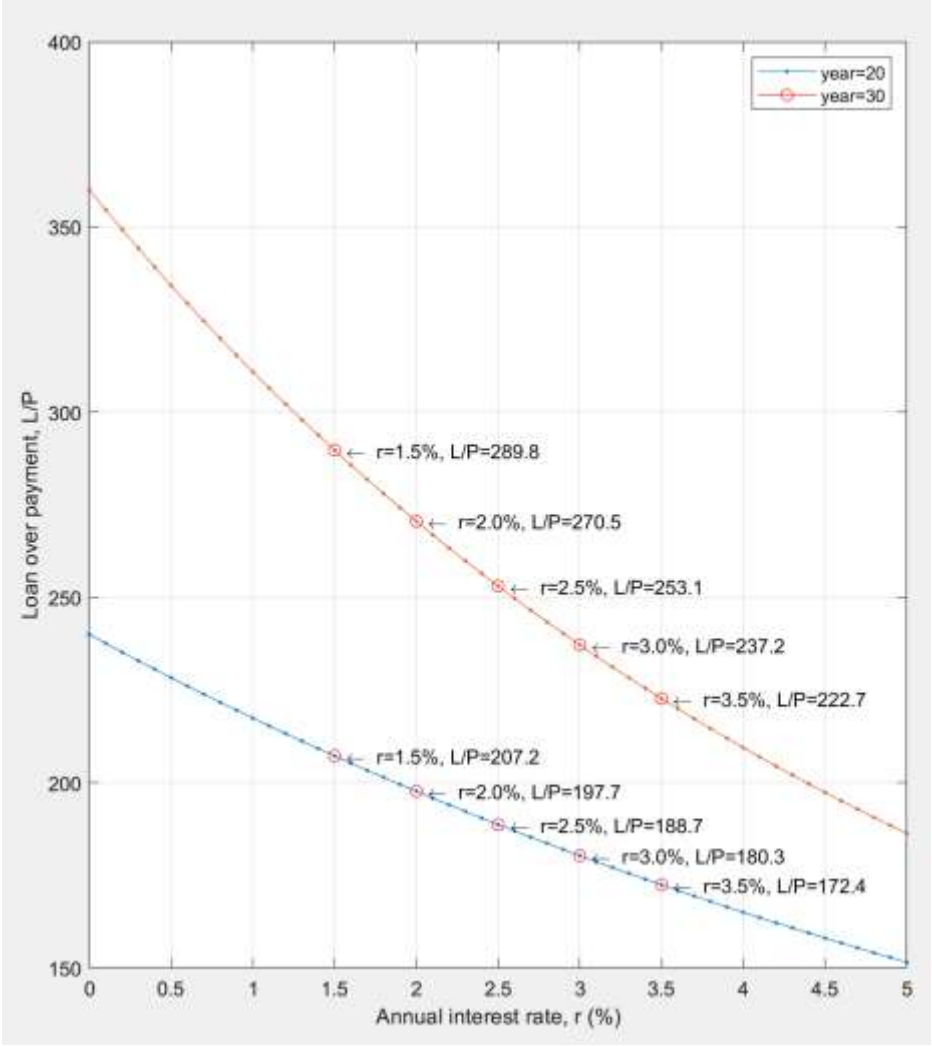
$$f_1 = f_2 \Rightarrow x = \frac{pr(1+r/12)^{240}}{(1+r/12)^{240}-1}$$

Rule of 200 (Roger's Formula)

- Loan/payment vs. r

$$L\left(1+\frac{r}{12}\right)^{12*20} = P\frac{(1+r/12)^{240}-1}{r/12}$$
$$\Rightarrow \frac{L}{P} = \frac{(1+r/12)^{240}-1}{(r/12)(1+r/12)^{240}} = \frac{1-(1+r/12)^{-240}}{r/12}$$

- Rule of 200
  - 假設20年房貸，利率2%
    ➔ 每月應付款 = 房貸總額/200
  - 延伸：若是30年房貸
    ➔ 每月應付款 = 房貸總額/270



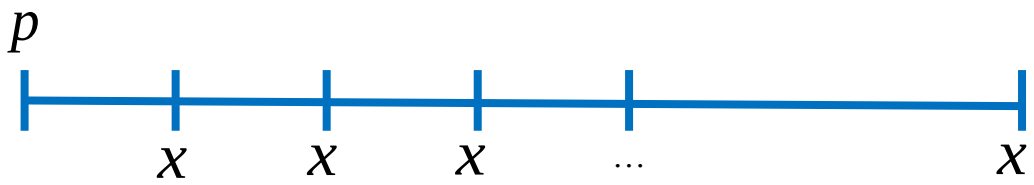
References

- References
  - Wiki: Time value of money
  - Wiki: Compound interest

Appendix

本息平均攤還：計算月付額的方法之二

- 現金流量圖



- Formula

Let b_i be the total unpaid amount at period i. Then
$b_0 = p$
$b_1 = p(1+r) - x$
$b_2 = (p(1+r) - x)(1+r) - x$
…
$\Rightarrow b_i = b_{i-1}(1+r) - x,$ with $b_0 = p$.

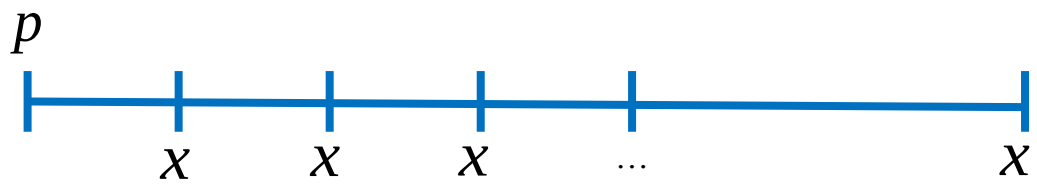
- Quiz

(a) Find the close−form solution of b_i.
(b) Set $b_n = 0$ to find the value of x.

hint: $b_i - x/r = (1+r)(b_{i-1} - x/r)$

本息平均攤還：計算月付額的方法之三

- 現金流量圖



- 本金及利息每期分攤

Let $x = p_i + q_i$,where p_i is the principal and $q_i$is the interest returned. Then

$q_1 = pr, p_1 = x - q_1$

$q_2 = (p - p_1)r, p_2 = x - q_2$

$q_3 = (p - p_1 - p_2)r, p_3 = x - q_3$

…

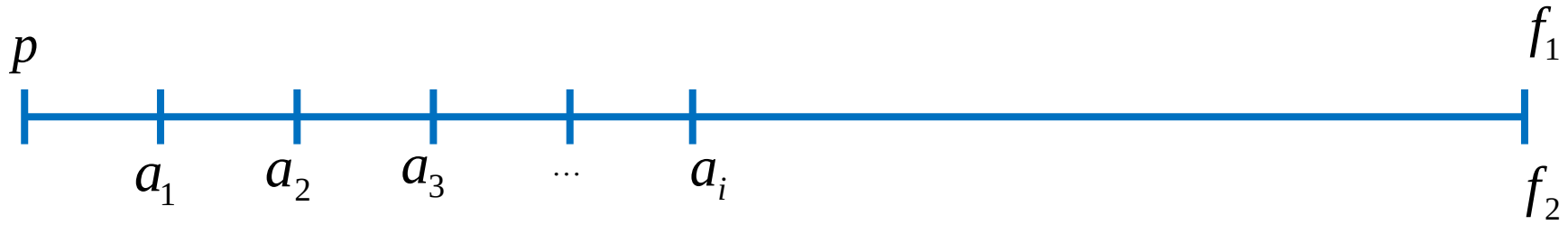
$$\Rightarrow q_i = \left(p - \sum_{k=1}^{i-1} p_k\right) r, p_i = x - q_i$$

- Quiz

(a) Find the recurrent formula of q_i.
(b) Find the close−form expression of q_i.

本金平均攤還法：計算月付額的方法之一

- 現金流量圖



$$
\begin{aligned}
a_1 &= p/n + pr \\
a_2 &= p/n + (p - p/n)r \\
a_3 &= p/n + (p - 2p/n)r \\
&\cdots \\
a_i &= p/n + p\left(1 - \frac{i-1}{n}\right)r
\end{aligned}
$$

Quiz:

$$
\begin{cases}
f_1 = p(1+r)^n \\
f_2 = \sum_{i=1}^{n} a_i (1+r)^{n-i}
\end{cases}
\Rightarrow \text{How to prove } f_1 = f_2 \text{ ?}
$$

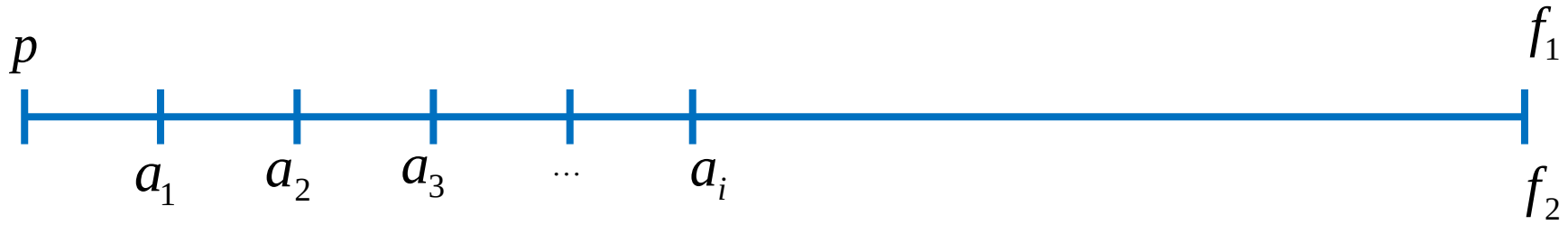
Hints:

$$S_0 = \sum (1+r)^{-i} = (1 - (1+r)^{-n})/r$$

$$S_1 = \sum i(1+r)^{-i} = (1+r)S_0/r - n(1+r)^{-n}/r$$

本金平均攤還法：計算月付額的方法之二

- 現金流量圖



Let $a_i = p_i + q_i$ be the periodic payment,
where p_i and q_i are principal and interest components, respectively. Then:

$p_1 = p_2 = \cdots = p_n = p/n$

$q_1 = pr$

$q_2 = (p - p/n)r$

$q_3 = (p - 2p/n)r$

$\cdots$

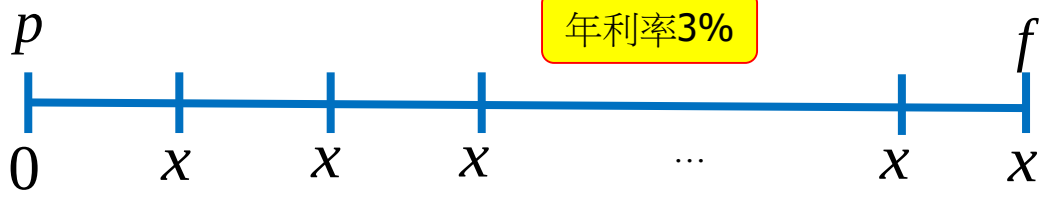
$q_i = (p - (i-1)p/n)r$

Useful Functions

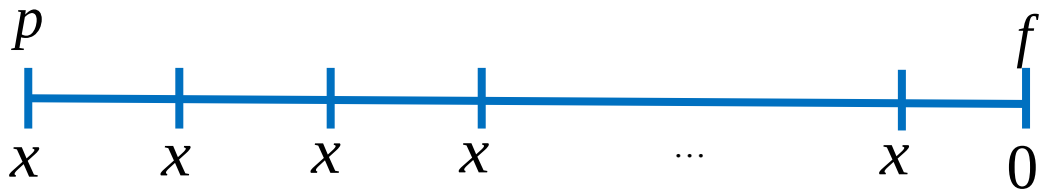
- Two MATLAB functions (utility toolbox) with self demo
  - loan.m
    - Monthly payment ➔ x=loan(p, 3/100, 20)

為期20年

年利率3%



  - saving.m
    - Present value ➔ p=saving(x, 3/100, 20, 'initial');
    - Final value ➔ f=saving(x, 3/100, 20, 'final');



Rule of 270

- 社會新鮮人的困惑
  - 若月收入3萬，房貸約佔1/3（1萬），為期30年 ➔ 台北市哪裡去找270萬的房子？
- 方法是人想出來的！
  - 從已無貸款的房子來進行增貸
    - 請和父母保持好關係，早晚問安、不時共餐

Internal Rate of Return (IRR)
內部報酬率

J.-S. Roger Jang (張智星)
MIR Lab, CSIE Dept.
National Taiwan University
jang@mirlab.org, http://mirlab.org/jang

2023/9/20

Internal Rate of Return (IRR)

- Definition
  - The rate at which an investment plan breaks even
- Also known as
  - Effective interest rate
  - Annualized effective compounded return rate
  - 年化報酬率
  - …
- Application
  - Comparison of 2 investment plans

Quiz:
Definition of IRR in plain texts!

The term “internal” indicates it does not take environmental factors (e.g., inflation) into consideration.

IRR Comparisons

- Is this investment plan good? Profit=250?

時間(年)	0	1	2	3	4	5
收支	-2000	450	450	450	450	450

  - Comparison with bank interest rate 2%

n年	0	1	2	3	4	5
收支	-2000	450	450	450	450	450
(1+Rate)^n	1.00	1.02	1.04	1.06	1.08	1.10
現值	-2000	441.18	432.53	424.05	415.73	407.58

NPV: Net present value

➔ NPV=121.07

  - Comparison with bank interest rate 5%

n年	0	1	2	3	4	5
收支	-2000	450	450	450	450	450
(1+Rate)^n	1.00	1.05	1.10	1.16	1.22	1.28
現值	-2000	428.57	408.16	388.73	370.22	352.59

➔ NPV=-51.73

- What is the equivalent interest rate?

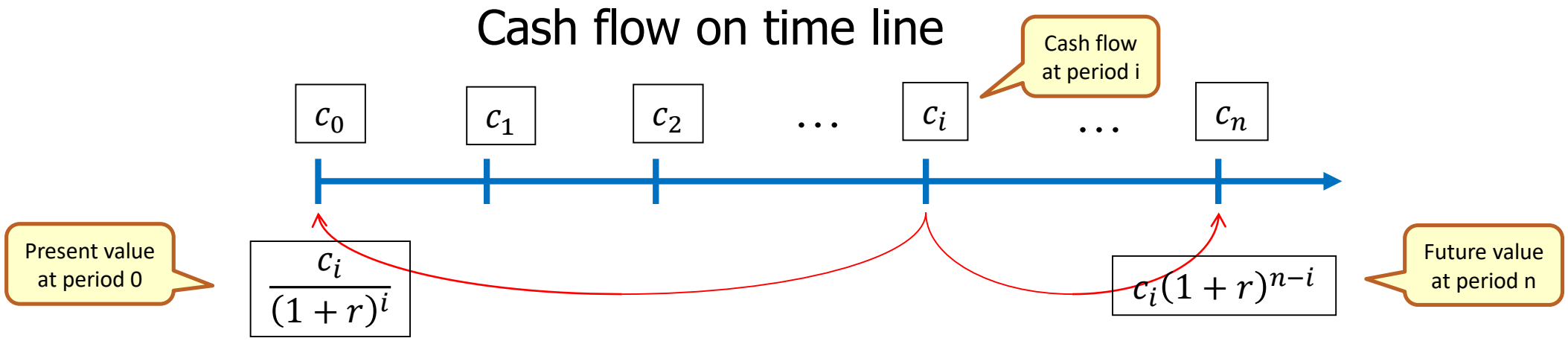
IRR Computation

- Net present value (NPV) = 0

$$NPV = \sum_{i=0}^{n} \frac{c_i}{(1+r)^i} = 0$$

- Net future value (NFV) = 0

$$NFV = \sum_{i=0}^{n} c_i(1+r)^{n-i} = 0$$



An IRR Example

- Cash flow table

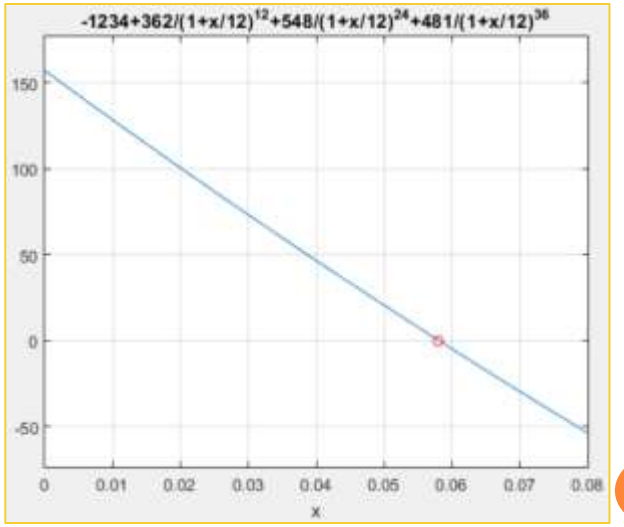
Year (i)	Cash flow (ci)
0	-1234
1	362
2	548
3	481

- Equations
  - Yearly compounding

$$NPV = -1234 + \frac{362}{(1+r)^1} + \frac{548}{(1+r)^2} + \frac{481}{(1+r)^3} = 0 \Rightarrow r = 5.96\%$$

  - Monthly compounding

$$NPV = -1234 + \frac{362}{(1+r/12)^{12}} + \frac{548}{(1+r/12)^{24}} + \frac{481}{(1+r/12)^{36}} = 0 \Rightarrow r = 5.80\%$$



儲蓄險比較

郵局六年期吉利保險

30歲男性繳 6 年：IRR 值 0.839%

6 年期儲蓄險 - 郵局六年期吉利保險			
年度	保費	領回	小計
0	-16320		-16320
1	-16157		-16157
2	-16157		-16157
3	-16157		-16157
4	-16157		-16157
5	-16157		-16157
6		100000	100000
		IRR	0.839%

Based on yearly compounding

遠雄好鑽養老保險

繳6年領回38萬：IRR 2.093%

六年期儲蓄險 - 遠雄好鑽養老 6 年期			
年度	保費	領回	小計
0	-59356		-59356
1	-58762		-58762
2	-58762		-58762
3	-58762		-58762
4	-58762		-58762
5	-58762		-58762
6		380,000	380,000
		IRR	2.093%

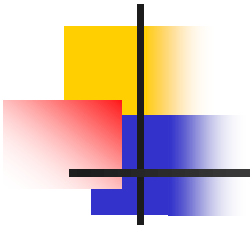
Based on yearly compounding

投資方案比較

- 三種投資方案
  - 方案一： 2年賺20% ➔ irr=9.54% (年複利) or 9.15% (月複利)
  - 方案二： 5年賺50% ➔ irr=8.45% (年複利) or 8.14% (月複利)
  - 方案三： 10年賺100% ➔ irr=7.18% (年複利) or 6.95% (月複利)
- Exercise
  - Write a Python function to execute the above computation.

References

- IRR
  - 綠角財經筆記 ← A very good example of IRR
  - Wiki
- 儲蓄險比較
  - IRR值EXCEL教學：儲蓄險比較利潤請看IRR值算年利率，而非報酬率
  - 六年儲蓄險台幣IRR分析-郵局六年期吉利保險(2012年)附DM
  - 六年儲蓄險台幣 IRR 分析 -遠雄好鑽養老保險 (2012附DM)



Matrix Formulas

張智星 (Roger Jang)
jang@mirlab.org
http://mirlab.org/jang
台大資工系 多媒體檢索實驗室

Transpose and Inverse

- Matrix transpose

$$(AB)^T = B^T A^T$$

$$(ABC)^T = C^T B^T A^T$$

$$\mathbf{a}^T\mathbf{b} = \mathbf{b}^T\mathbf{a}$$

$$\|\mathbf{x}\|^2 = \mathbf{x}^T\mathbf{x}$$

- Matrix inverse

$$(AB)^{-1} = B^{-1}A^{-1}$$

$$(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$$

Block Form of a Matrix (1/2)

- Matrix partition into a block form:
  - Examples

$$A = \left[\begin{array}{c|cc} 1 & 2 & 3 \\ \hline 4 & 5 & 6 \\ \hline 7 & 8 & 9 \end{array}\right] = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

$$A = \begin{bmatrix} | & & | & & | \\ \mathbf{a}_1 & \cdots & \mathbf{a}_j & \cdots & \mathbf{a}_n \\ | & & | & & | \end{bmatrix} = \begin{bmatrix} - & \mathbf{b}_1^T & - \\ & \vdots & \\ - & \mathbf{b}_i^T & - \\ & \vdots & \\ - & \mathbf{b}_m^T & - \end{bmatrix}$$

Column vector!

Row vector!

Block Form of a Matrix (2/2)

- Block-form matrix operations
  - Examples

$$A=\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, B=\begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

$$\Rightarrow A^T=\begin{bmatrix} A_{11}^T & A_{21}^T \\ A_{12}^T & A_{22}^T \end{bmatrix}, AB=\begin{bmatrix} A_{11}B_{11}+A_{12}B_{21} & A_{11}B_{12}+A_{12}B_{22} \\ A_{21}B_{11}+A_{22}B_{21} & A_{21}B_{12}+A_{22}B_{22} \end{bmatrix}$$

$$A=\begin{bmatrix} | & | & | \\ \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \\ | & | & | \end{bmatrix}, x=\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \Rightarrow A\mathbf{x}=\mathbf{a}_1x_1+\mathbf{a}_2x_2+\mathbf{a}_3x_3$$

Gradient of a Function

- Gradient of a function $f(\mathbf{x})$

Quiz!

$$\nabla f(\mathbf{x}) = \begin{bmatrix} \partial f(\mathbf{x})/\partial x_1 \\ \vdots \\ \partial f(\mathbf{x})/\partial x_n \end{bmatrix}, \text{ where } \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

- If $f(\mathbf{x}) = \mathbf{c}^T\mathbf{x} = \mathbf{x}^T\mathbf{c}$

$$\nabla f(\mathbf{x}) = \mathbf{c} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

Quadratic Form (1)

- Quadratic form of $\mathbf{x}$

Quiz!

$$\mathbf{x}^T A\mathbf{x} = \sum_{i=1}^{n} a_{ii} x_i^2 + \sum_{i=1}^{n} \sum_{j=1, j\neq i}^{n} a_{ij} x_i x_j$$

- A can be assumed symmetric since

$$\mathbf{x}^T A\mathbf{x} = \mathbf{x}^T \frac{A + A^T}{2} \mathbf{x}$$

Quadratic Form (2)

- When $\mathbf{x} = [x, y]^T$

Quiz!

$$\mathbf{x}^T A\mathbf{x} = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}x + a_{21}y & a_{12}x + a_{22}y \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$= a_{11}x^2 + (a_{12} + a_{21})xy + a_{22}y^2$$

- Different values of A can lead to the same quadratic form as long as $(a_{12} + a_{21})$ is the same.

Quadratic Form (3)

- When $\mathbf{x} = [x, y, z]^T$

Quiz!

$$\mathbf{x}^T A\mathbf{x} = \begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Expand it by Brutal force!

$$= a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + (a_{12} + a_{21})xy + (a_{13} + a_{31})xz + (a_{23} + a_{32})yz$$

- Different values of A can lead to the same quadratic form too.

Gradient of a Quadratic Form

- The gradient of a quadratic form

$$\nabla\left(\mathbf{x}^T A\mathbf{x}\right) = \begin{cases} 2A\mathbf{x}, \textit{if } A \textit{ is symetric} \\ \left(A + A^T\right)\mathbf{x}, \textit{otherwise} \end{cases}$$

Quiz!

- Example

$$\mathbf{x}^T A\mathbf{x} = \begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

$$= a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + \left(a_{12} + a_{21}\right)xy + \left(a_{13} + a_{31}\right)xz + \left(a_{23} + a_{32}\right)yz$$

$$\nabla\left(\mathbf{x}^T A\mathbf{x}\right) = \begin{bmatrix} 2a_{11}x + \left(a_{12} + a_{21}\right)y + \left(a_{13} + a_{31}\right)z \\ 2a_{22}y + \left(a_{12} + a_{21}\right)x + \left(a_{23} + a_{32}\right)z \\ 2a_{33}z + \left(a_{13} + a_{31}\right)x + \left(a_{23} + a_{32}\right)y \end{bmatrix}$$

$$\Rightarrow = \begin{bmatrix} 2a_{11} & a_{12} + a_{21} & a_{13} + a_{31} \\ a_{12} + a_{21} & 2a_{22} & a_{23} + a_{32} \\ a_{13} + a_{31} & a_{23} + a_{32} & 2a_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

$$= \left(A + A^T\right)\mathbf{x}$$

Common Formulas

- Some common formulas (assuming A is symmetric and all derivatives are w.r.t x)

$$\nabla\left(\mathbf{x}^T\mathbf{c}\right)=\nabla\left(\mathbf{c}^T\mathbf{x}\right)=\mathbf{c}$$

$$\nabla\left(\mathbf{x}^T\mathbf{x}\right)=2\mathbf{x}$$

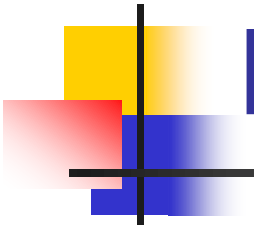
Quiz!

$$\nabla\left(\mathbf{x}^T A\mathbf{c}\right)=A\mathbf{c}$$

$$\nabla\left(\mathbf{c}^T A\mathbf{x}\right)=A^T\mathbf{c}$$

$$\nabla\left(\mathbf{x}^T A\mathbf{x}\right)=2A\mathbf{x}$$

$$\nabla\left(\mathbf{x}^T A\mathbf{x}+\mathbf{b}^T\mathbf{x}+\mathbf{c}\right)=2A\mathbf{x}+\mathbf{b}$$



Reference

- Matrix cookbook
  - http://www2.imm.dtu.dk/pubdb/views/edoc_download.php/3274/pdf/imm3274.pdf

Intro to Machine Learning

J.-S. Roger Jang (張智星)

jang@mirlab.org

http://mirlab.org/jang

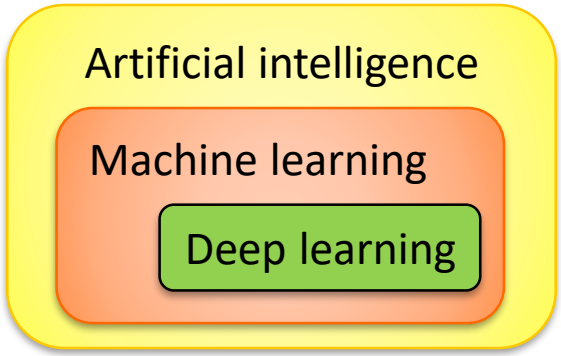
MIR Lab, CSIE Dept.

National Taiwan University

2023/9/20

Machine Learning (ML)

- Definition
  - Field of study that gives computers the ability to learn without being explicitly programmed – by Arthur Samuel, 1959.
  - Computational methods that use existing data to make predictions
- Concepts
  - 視其所以，觀其所由，察其所安，人焉廋哉？人焉廋哉？
  - 近朱者赤、近墨者黑



More about Machine Learning

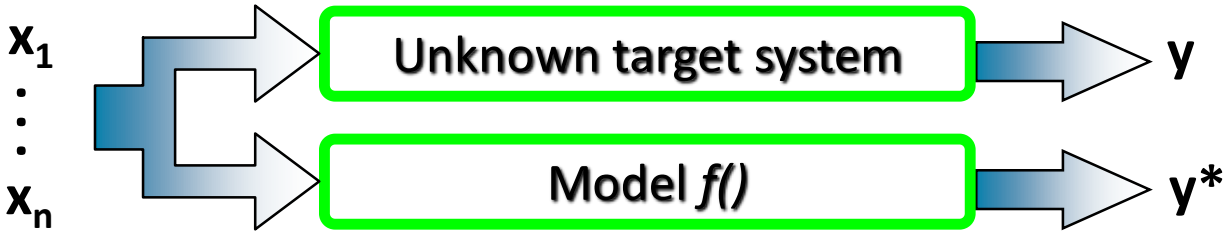
- Quick examples
  - Use height and weight to predict gender
  - Use historical data to predict stock market
- Comparison
  - Before ML ➔ Rule-based
  - After ML ➔ Data-driven
- Prerequisites for ML
  - Probability & statistics
  - Linear algebra
  - Optimization
  - …

Broad Areas of ML

- Classification: Assign a category to a given object
  - Determine if it will rain tomorrow
- Regression: Predict a real-value for a given object
  - Predict the exchange rate
- Ranking: Order objects according to some criterion
  - Rank the webpages returned by a search engine
- Clustering: Partition data into homogeneous groups
- Dimensionality reduction: Find low-dimensional manifold preserving some properties of the data
- Density estimation: Learning probability density function according to sample data

Concept of Modeling

- Given desired i/o pairs (training set) of the form (x_1, ..., x_n; y), construct a model to match the i/o pairs



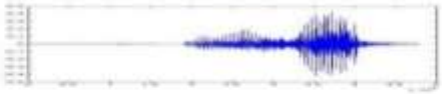
- Two basic steps in modeling

Quiz!

  - structure identification: input selection, model complexity
    - Example: order determination in polynomial fitting
  - parameter identification: optimal parameters
    - Example: coefficient determination in polynomial fitting

Example of Machine Learning (or Modeling)

- Speech Recognition

$f^*($  $) =$ "Morning"

- Handwritten Recognition

$f^*($  $) =$ "2"

- Playing Go

$f^*($  $) =$ "5-5" (step)

- Dialogue System

$f^*($ "Hi" $) =$ "Hello"

(what the user said) (system response)

https://www.slideshare.net/tw_dsconf/ss-70083878

Three Basic Paradigms of ML

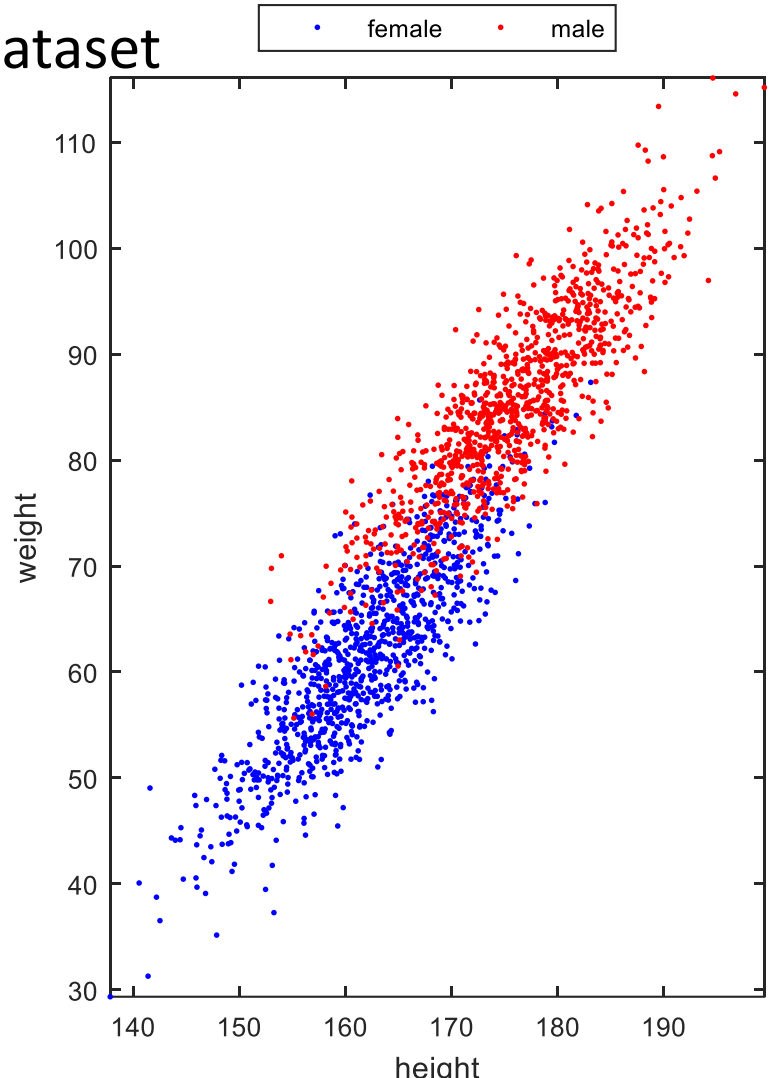
Quiz!

- Supervised learning
  - Each input has output, which is available immediately or with a delay of fixed duration.
  - Two types: classification & regression
  - Examples: Weather prediction, face recognition, …
- Reinforcement learning
  - The output is available but with a delay of variable duration.
  - Example: Chess playing, control system, path planning…
- Unsupervised learning
  - No output part is involved.
  - Example: K-means clustering, principal component analysis, …

GHW Dataset

- GHW (gender-height-weight) dataset
  - 10000 entries with 3 columns
    - Gender: Categorical
    - Height: Numerical
    - Weight: Numerical
  - Source

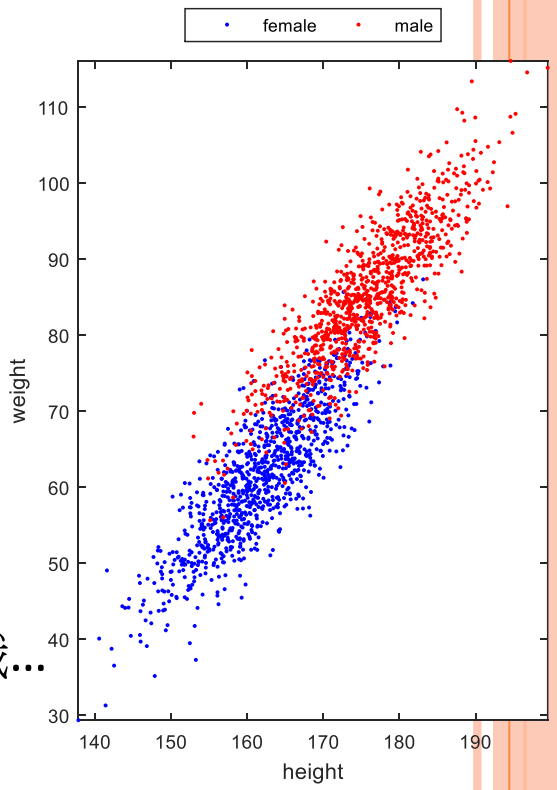
	A	B	C	D	E
1	Gender	Height	Weight	Height (cm)	Weight (kg)
2	Male	73.847017	241.89356	187.5714232	109.72099
3	Male	68.781904	162.31047	174.7060363	73.622732
4	Male	74.110105	212.74086	188.2396677	96.49755
5	Male	71.730978	220.04247	182.1966851	99.809504
6	Male	69.881796	206.3498	177.4997615	93.598619
7	Male	67.253016	152.21216	170.8226598	69.042216
8	Male	68.785081	183.92789	174.7141064	83.428219
9	Male	68.348516	167.97111	173.6052294	76.190352
10	Male	67.01895	175.92944	170.2281321	79.800187
11	Male	63.456494	156.39968	161.1794947	70.941642
12	Male	71.195382	186.60493	180.836271	84.642501
13	Male	71.640805	213.74117	181.967645	96.951285
14	Male	64.766329	167.12746	164.506476	75.807679
15	Male	69.28307	189.44618	175.978098	85.931272



Multiple Ways of Prediction

- Multiple ways of prediction based on GHW dataset
  - Classification: Height, weight ➔ gender
  - Regression: Gender, height ➔ weight
  - Regression: Gender, weight ➔ height
- In general
  - Classification: the output is categorical
  - Regression: the output is numerical
- Sometimes it's not so obvious!
- 範例
  - 分類：信用卡盜刷、警示帳戶、支票金額辨識...
  - 迴歸：房屋估價、貸款金額、收入預估...

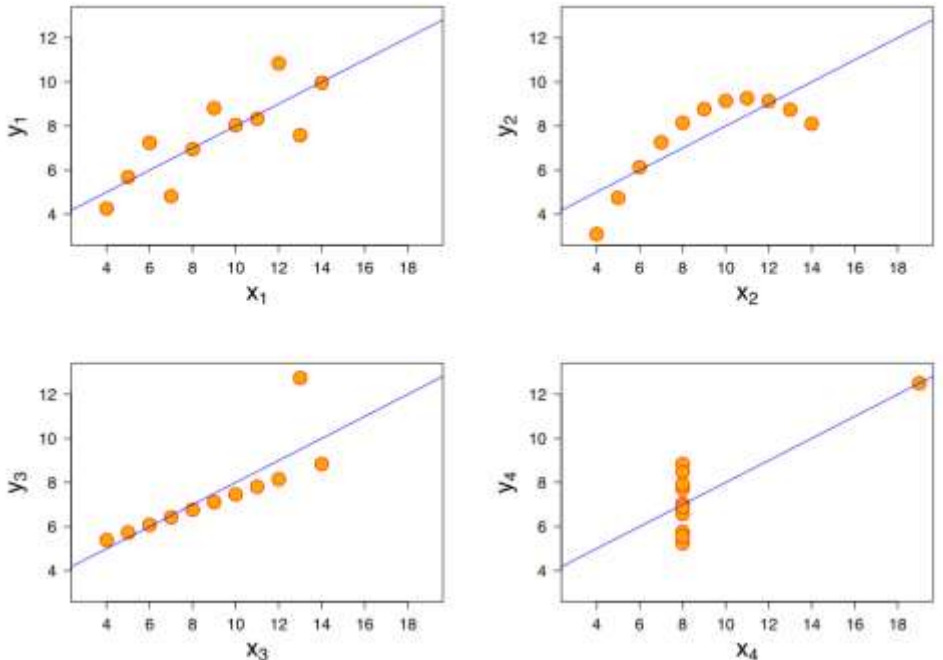
Quiz!



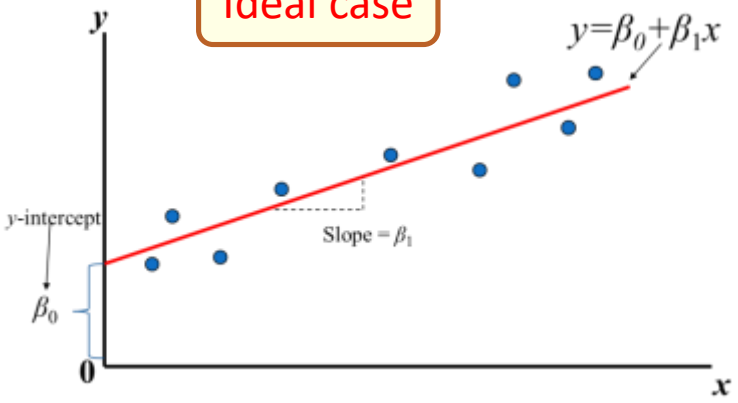
Basic Case of Linear Regression: Polynomial Fitting

- Structure determination
  - $y = \beta_0 + \beta_1 x$
  - $y = \beta_0 + \beta_1 x + \beta_2 x^2$
  - $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3$
- Parameter identification
  - Least-squares method

Real cases

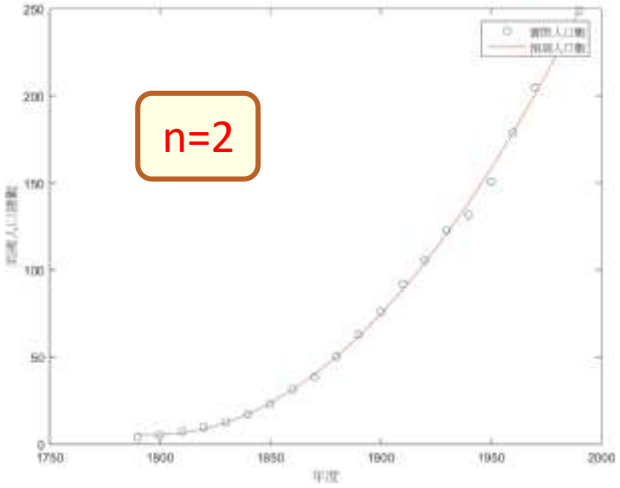


Ideal case

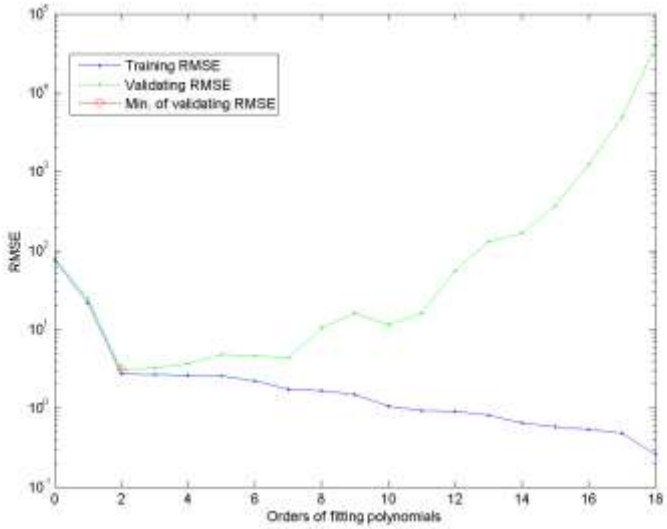


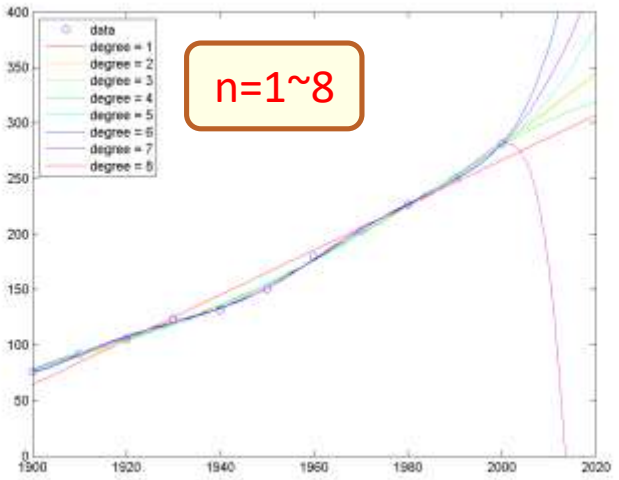
Example of Polynomial Fitting

- 美國人口總數預測
  - 模型
    - 多項式 $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \cdots + \beta_n x^n$
  - 模型複雜度：冪次 n
  - 模型效能評估
    - 交叉驗證 (cross validation)



Cross validation





More ML Application Examples

- Optical character recognition
- Document classification
- Part-of-speech tagging
- Speech processing
  - Recognition
  - synthesis
- Image recognition
  - Face recognition
- Info. retrieval
  - Recommendation
  - search
- Security
  - Fraud detection (credit card, telephone)
  - Network intrusion
  - Video surveillance
  - Speaker id.
- Self-driving cars

What People Do with ML

- For practitioners
  - Acquire application domain knowledge
  - Know properties of machine learning methods
  - Get familiar with ML tools
- For researchers
  - Design new learning algorithms
  - Speed up the learning process
  - Derive theoretical bound of accuracy
  - Take care of data
    - Big data
    - Imbalanced data
    - Missing data
    - …

List of Classifiers

- Commonly used classifiers
  - K-nearest-neighbor classifiers
  - Quadratic classifiers
    - Naïve Bayes classifiers
  - Linear classifiers
    - Single-layer perceptrons
    - SVM
  - Neural networks
    - Multilayer perceptrons
      - DNN
    - Radial-basis function networks
  - Classification and regression trees (CART)
    - Random forests
  - Many many more…

Commonly Used Datasets for ML

J.-S. Roger Jang (張智星)
jang@mirlab.org
http://mirlab.org/jang
MIR Lab, CSIE Dept.
National Taiwan University

Datasets

- Numerous datasets for testing ML algorithms
  - Kaggle
  - UCI Machine Learning Repository
  - Image net
  - MNIST handwritten digit database
  - Labeled Faces in the Wild
  - Many many more…
- Looking for a specific dataset?
  - Google search engine
  - Google dataset search

Try "license plate dataset" in Google!

UCI Dataset: Iris

- Source
  - R.A. Fisher, 1936
- Goal
  - Predict the types of iris in Hawaii
- Dataset specs
  - 150 instances, 3 classes
  - 4 attributes (features)
    - sepal length
    - sepal width
    - petal length
    - petal width

UCI Dataset: Wine

- Source
  - Institute of Pharmaceutical and Food Analysis and Technologies, Via Brigata Salerno, 16147 Genoa, Italy.
- Goal
  - Using 13 chemical constituents to determine the origin of wines
- Dataset specs
  - 178 instances, 3 classes, 13 attributes

UCI Dataset: Abalone

- Source
  - Dept. of Primary Industry and FIsheries, Tasmania, Australia
- Goal
  - Predict the age of abalone (鮑魚)
- Dataset specs
  - 4177 instances, 29 classes
  - 8 attributes (features): sex, length, diameter, height, whole weight, shucked weight, viscera weight, shell weight
  - 1 output: rings (+1.5 gives the age in years)

UCI Dataset: Mushroom Classification

- Source
  - Mushroom records drawn from The Audubon Society Field Guide to North American Mushrooms (1981)
- Goal
  - To determine a mushroom is poisonous or edible
- Dataset specs
  - 8124 instances, 2 classes, 22 attributes

UCI Dataset: Liver Disorder

- Source
  - BUPA Medical Research Ltd.
- Goal
  - Use variables from blood tests and alcohol consumption to see if liver disorder exists
- Dataset specs
  - 345 instances, 2 classes, 6 attributes (the first five are results from blood tests, the last one is alcohol consumption per day)

UCI Dataset: Credit Screening

- Source
  - Chiharu Sano, csano@bonnie.ICS.UCI.EDU
- Goal
  - Determine people who are granted credit
- Dataset specs
  - 125 instances, 2 classes, 15 attributes

UCI Dataset: House Price Prediction

- Source
  - CMU StatLib Library
- Goal
  - Predict house price near Boston
- Dataset specs
  - 506 instances, 13 attributes

1. CRIM: per capita crime rate by town
2. ZN: proportion of residential land zoned for lots over 25,000 sq.ft.
3. INDUS: proportion of non-retail business acres per town
4. CHAS: Charles River dummy variable (= 1 if tract bounds river; 0 otherwise)
5. NOX: nitric oxides concentration (parts per 10 million)
6. RM: average number of rooms per dwelling
7. AGE: proportion of owner-occupied units built prior to 1940
8. DIS: weighted distances to five Boston employment centres
9. RAD: index of accessibility to radial highways
10. TAX: full-value property-tax rate per $10,000
11. PTRATIO: pupil-teacher ratio by town
12. B: 1000(Bk - 0.63)^2 where Bk is the proportion of blacks by town
13. LSTAT: % lower status of the population
14. MEDV: Median value of owner-occupied homes in $1000's

MNIST Digit Dataset (1/2)

- Source

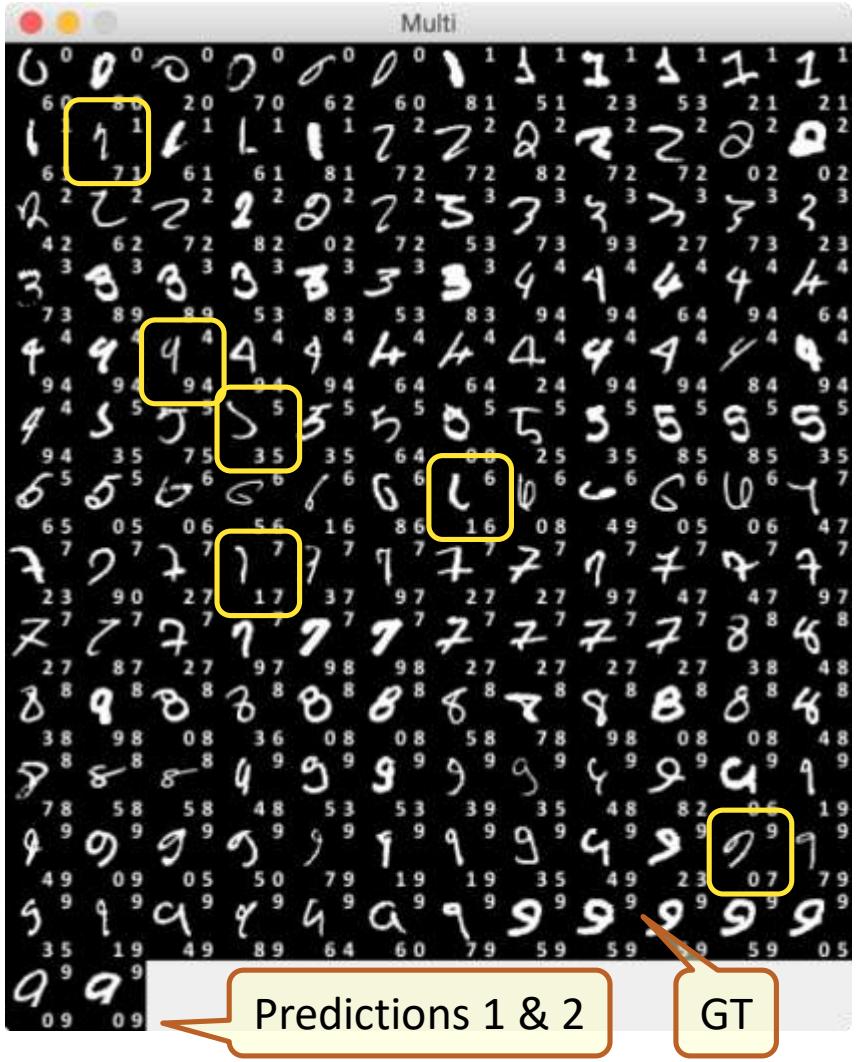
  Quiz: Full name of NIST?

  - NIST's Special Database 3 (collected among Census Bureau employees) and Special Database 1 (collected among high-school students)
- Goal
  - Recognize isolated hand-written digits of 0-9
- Dataset specs
  - 70000 instances
    - 60000 for training (30000 from SD-3 and 30000 from SD-1) of about 250 writers
    - 10000 for test (5000 from SD-3 and 5000 from SD-1)

    Disjoint writers!
  - Normalized to 28x28 gray-scale image, centered by gravity

MNIST Digit Dataset (2/2)

- Links
  - Data source
  - Wikipedia
- Examples



How to Acquire/Visualize the Datasets?

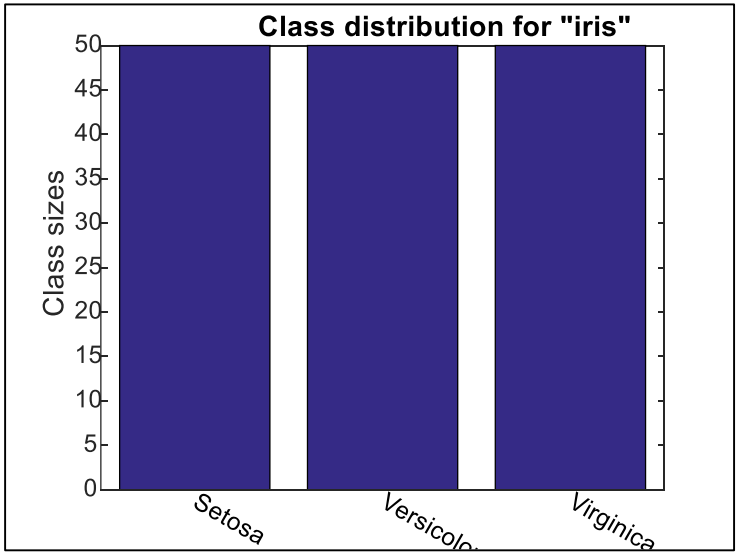
- Acquire the datasets
  - prData.m for acquiring PR data
  - dcData.m for acquiring DC data
- Visualize the datasets
  - Please refer to Chapter 2 of DCPR tutorial
- Example:

You need to download Machine Learning Toolbox to try these commands.

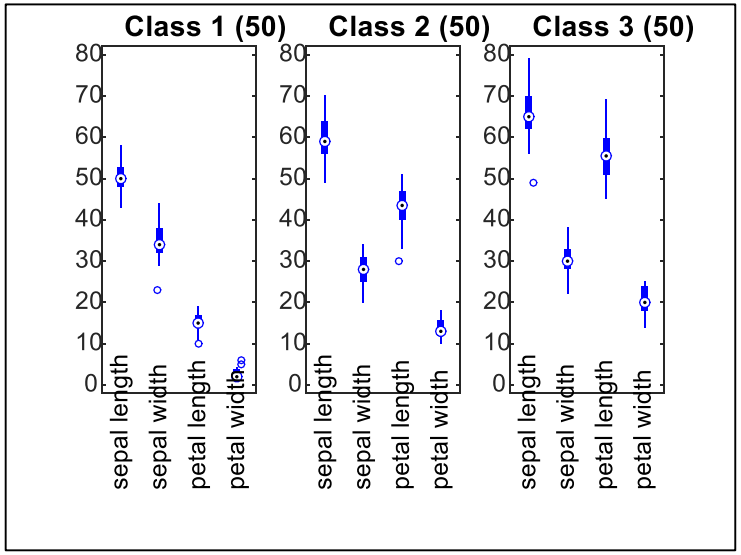
```
>> ds=prData('iris')

ds =
     dataName: 'iris'
    inputName: {'sepal length'  'sepal width'  'petal length'  'petal width'}
   outputName: {'Setosa'  'Versicolour'  'Virginica'}
        input: [4x150 double]
       output: [1x150 double]
```

Iris Dataset Visualization (1/2)

```
ds=prData('iris');
classSize=dsClassSize(DS, 1);
```

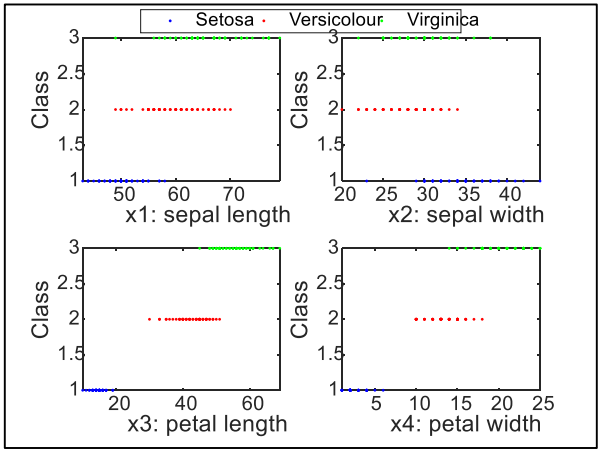



```
ds=prData('iris');
dsDistPlot(ds);
```

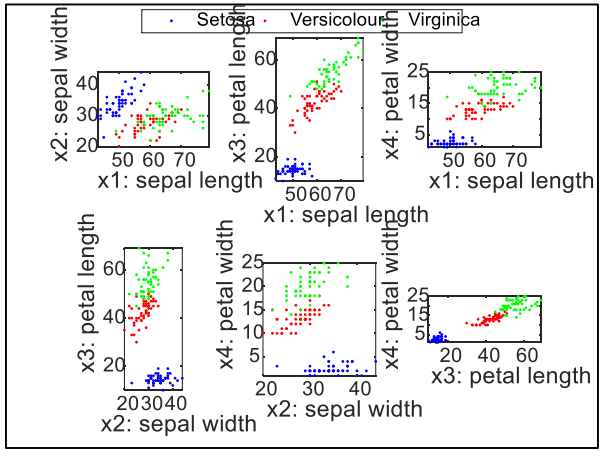


Iris Dataset Visualization (2/2)

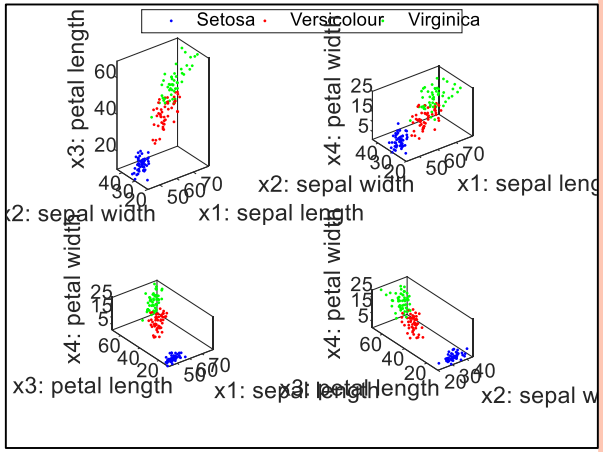
```
ds = prData('iris');
dsProjPlot1(ds);
```




```
ds = prData('iris');
dsProjPlot2(ds);
```




```
ds = prData('iris');
dsProjPlot3(ds);
```



K-Nearest Neighbor Classifiers (KNNC)

J.-S. Roger Jang (張智星)
jang@mirlab.org
http://mirlab.org/jang
MIR Lab, CSIE Dept.
National Taiwan University

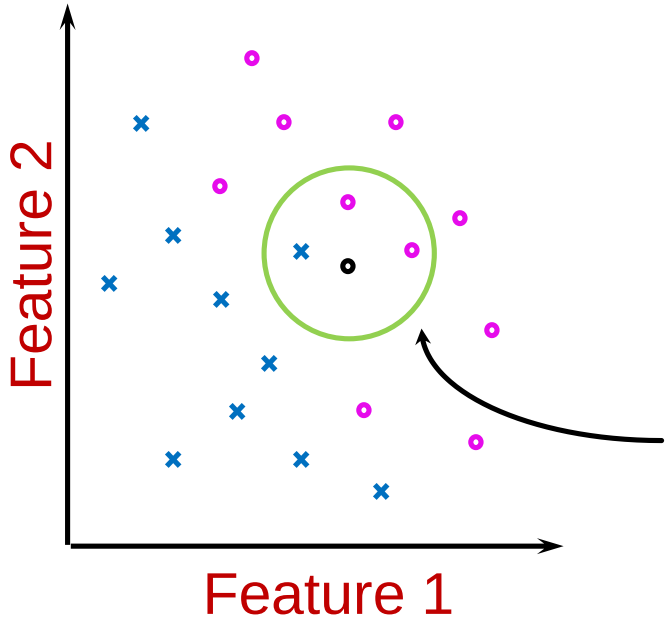
2023/9/20

Concept of KNNC

- Concept: 近朱者赤、近墨者黑
- Two Steps:

Quiz!

  - Find the first k nearest neighbors of a given point.
  - Determine the class of the given point by voting among k nearest neighbors.

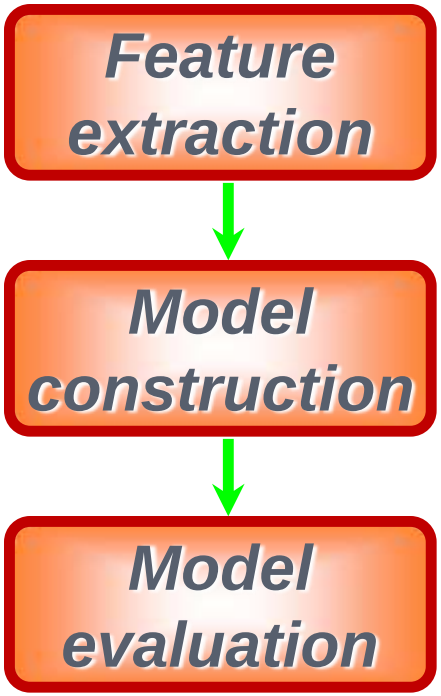


× : class-A point
∘ : class-B point
∘ : point with unknown class

3 nearest neighbors
➔ The point is classified as B via 3NNC.

Flowchart for KNNC

Flowchart of classification:



KNNC:

From raw data to features

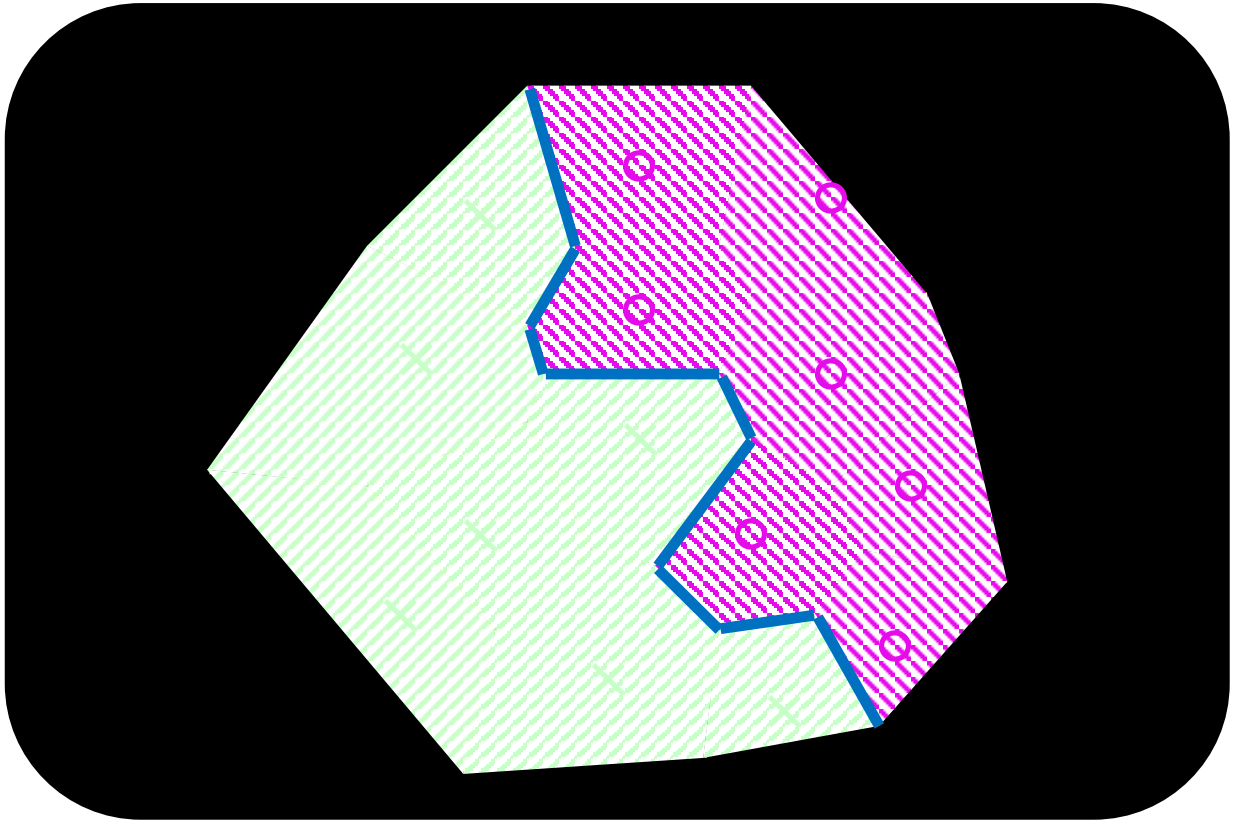
Clustering (optional)

KNNC evaluation on test dataset

Decision Boundary for 1NNC

- Voronoi diagram: piecewise linear boundary

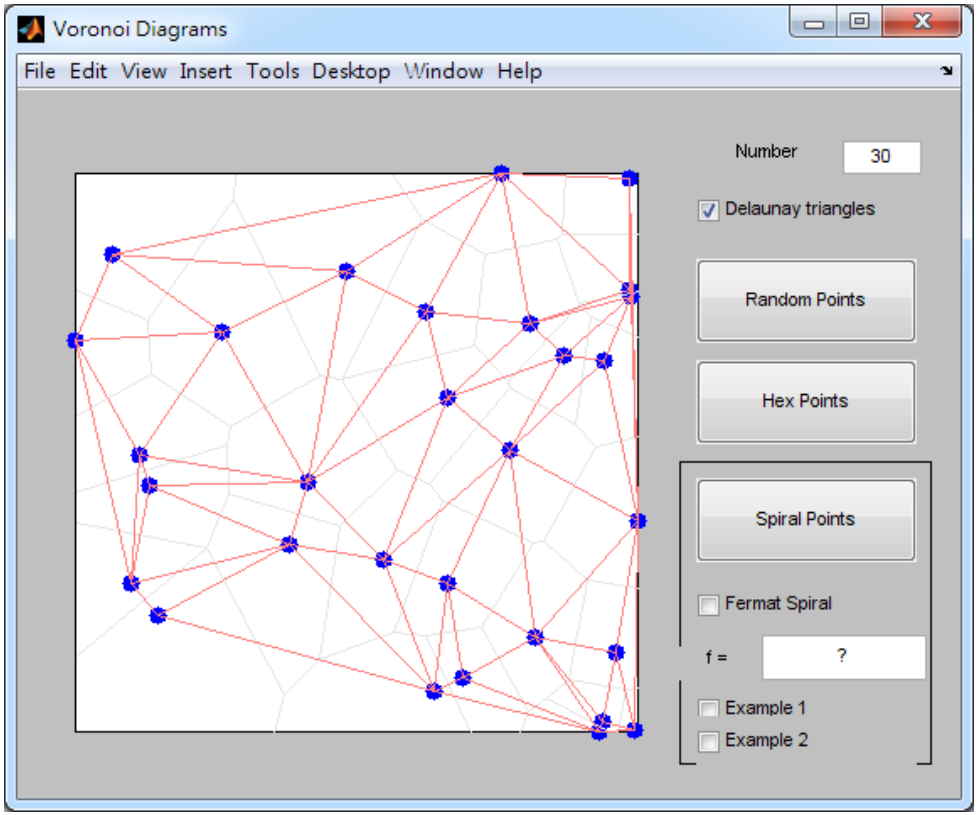
Quiz!



More about Voronoi diagrams

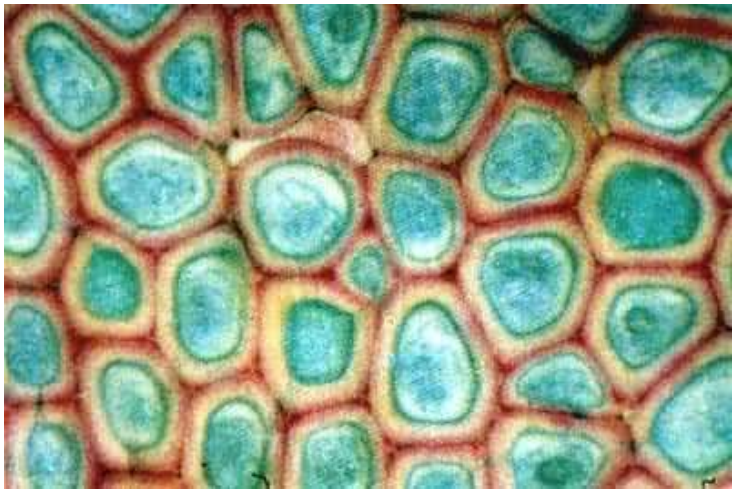
Demos by Cleve Moler

- Cleve's Demos of Delaunay triangles and Voronoi diagram
  - books/dcpr/example/cleve/vshow.m



Natural Examples of Voronoi Diagrams (1/2)

Natural Examples of Voronoi Diagrams (2/2)



Characteristics of KNNC

- Strengths of KNNC

  Quiz!

  - Intuitive
  - No computation for model construction
- Weakness of KNNC
  - Massive computation required when dataset is big
  - No straightforward way
    - To determine the value of K
    - To rescale the dataset along each dimension

Preprocessing of Feature Normalization

- Z normalization or z score
  - To have zero mean and unit variance along each feature
- Range normalization
  - To have a specific range, such as [0, 1], along each feature

Quiz!

Let $\mathbf{x} = [x_1, x_2, \cdots, x_n]$ be the values of a specific feature of a dataset

Z normalization:

$$\hat{x}_i = \frac{x_i - \mu}{\sigma}, \text{ with } \mu \text{ and } \sigma^2 \text{ being the sample mean and sample variance of } \mathbf{x} \text{ respectively}$$

Range normalization:

$$\hat{x}_i = \frac{x_i - \min(\mathbf{x})}{\max(\mathbf{x}) - \min(\mathbf{x})} \text{ to have a range of } [0, 1]$$

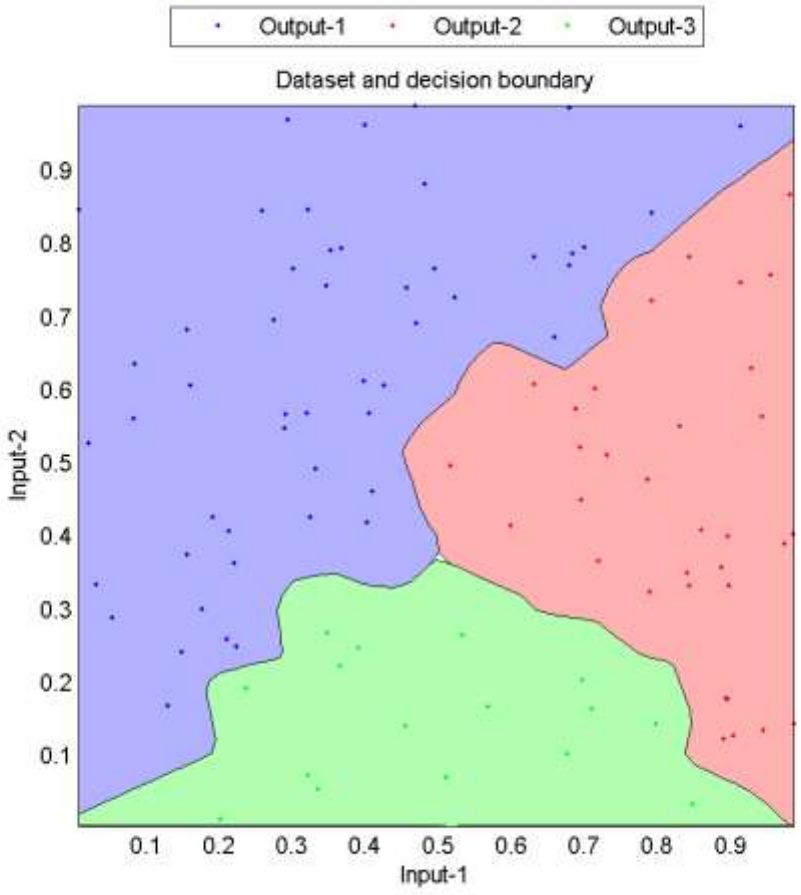
Variants for KNNC

- Many variants of KNNC:
  - Nearest prototype classification
    - Single prototype for each class ➔ Use “mean” or “average”
    - Several prototypes for each class ➔ Use “k-means clustering”
  - Distance-weighted votes
  - Edited nearest neighbor classification
  - k+k-nearest neighbor

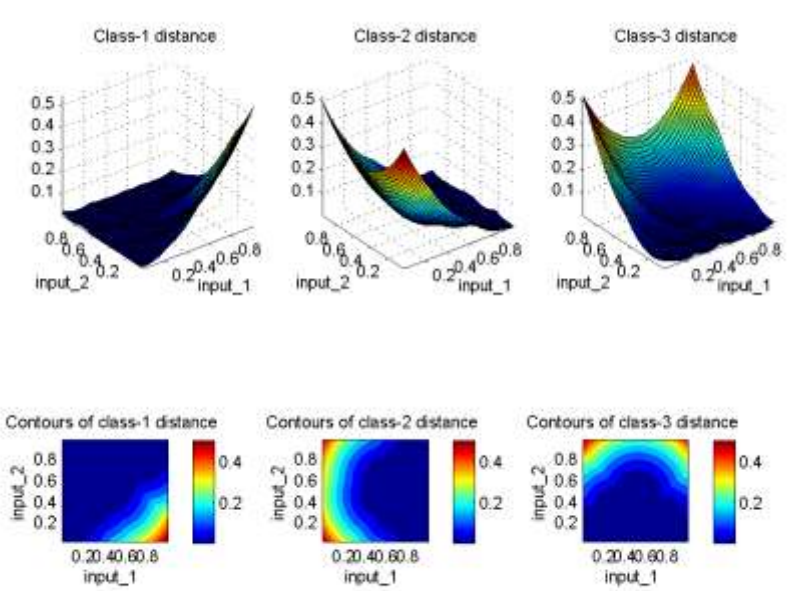
Quiz!

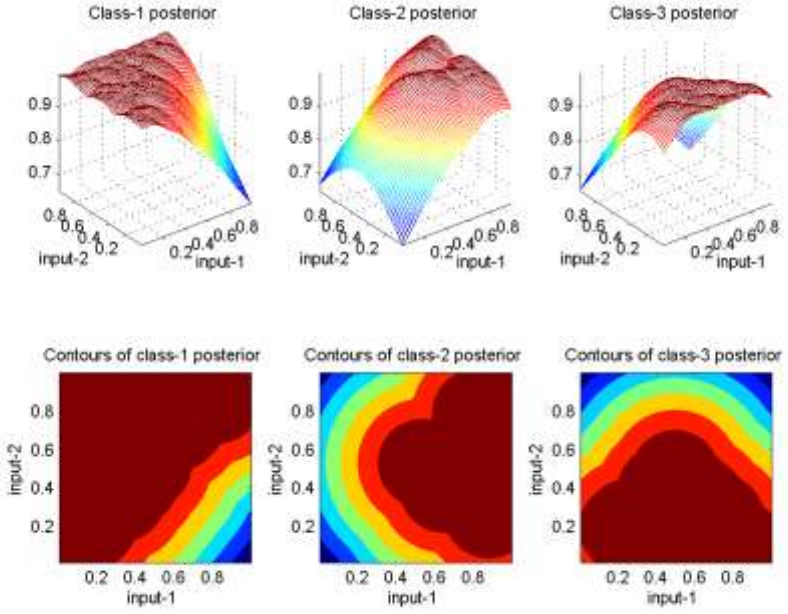
1NNC Decision Boundaries

- 1NNC Decision boundaries



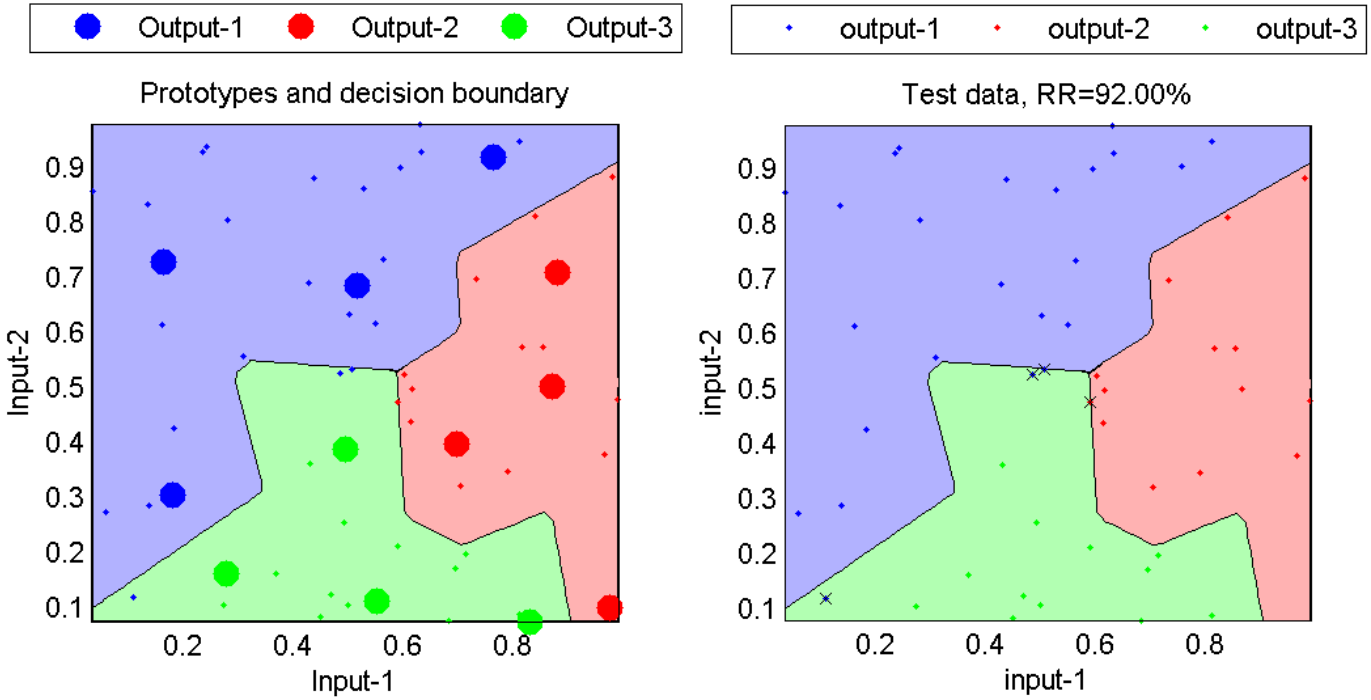
1NNC Distance/Posterior as Surfaces and Contours





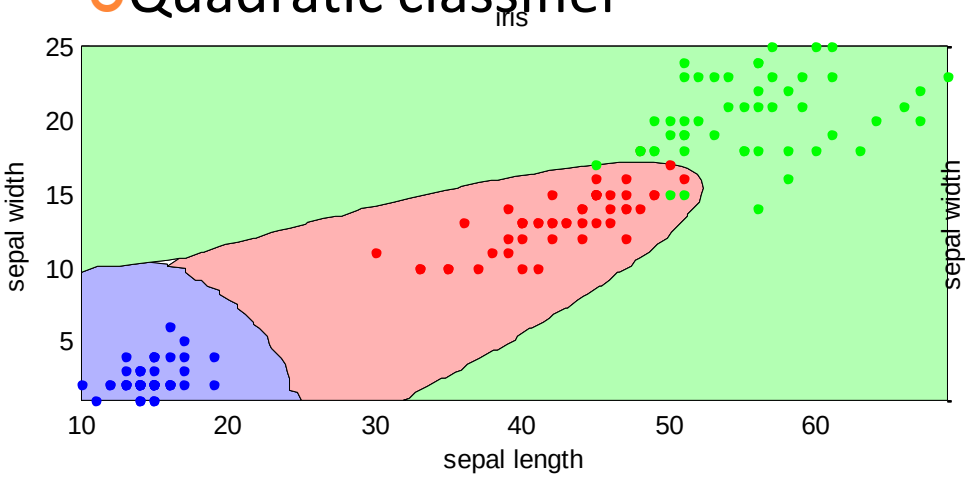
Using Prototypes in KNNC

- No. of prototypes for each class is 4.

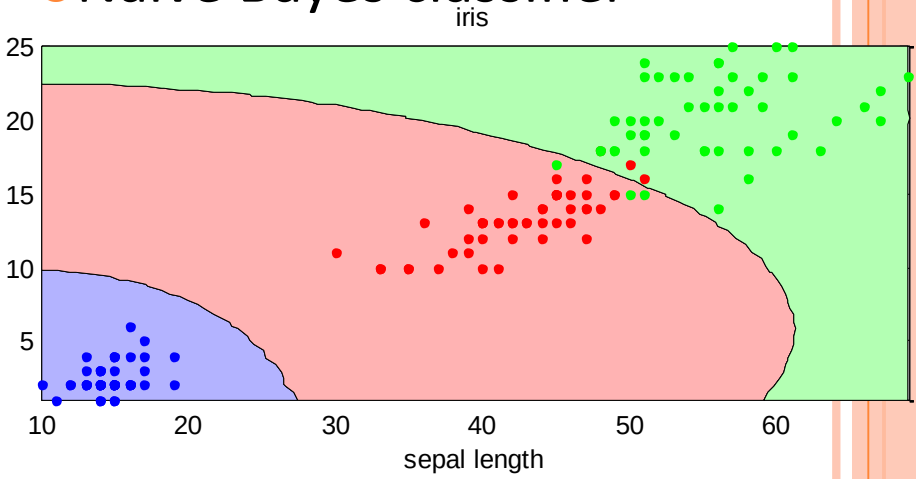


Decision Boundaries of Different Classifiers

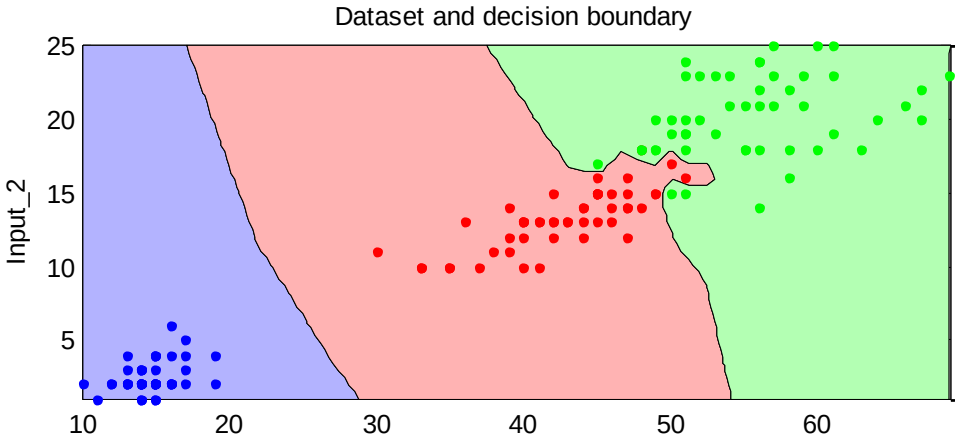
- Quadratic classifier



- Naive Bayes classifier



- 1NNC classifier



Exercise: KNNC Decision Boundary

- Given 6 samples of two classes as shown below, plot the decision boundary based on KNNC with k=1.

Exercise: Nearest Prototype Classifier

- If we want to use the nearest prototype classifier instead of KNNC, what methods be used to find the prototype(s) for each class in the following two cases?
  - When the number of prototype is 1 for each class.
  - When the number of prototype is more than 1 for each class.