

105-1: EE4052
計算機程式設計
Computer Programming

Unit 05: 多維度資料格式

連 豐 力

臺大電機系

Sep 2016 - Jan 2017

簡介

計算機程式設計 - 2016F
Chap 05: 多維度資料
Feng-Li Lian @ NTU-EE

■ 一維的數據

10

70	84	39	32	60	31	78	77	61	79
----	----	----	----	----	----	----	----	----	----

70
84
39
32
60
31
78
77
61
79

■ 二維的數據

3

10

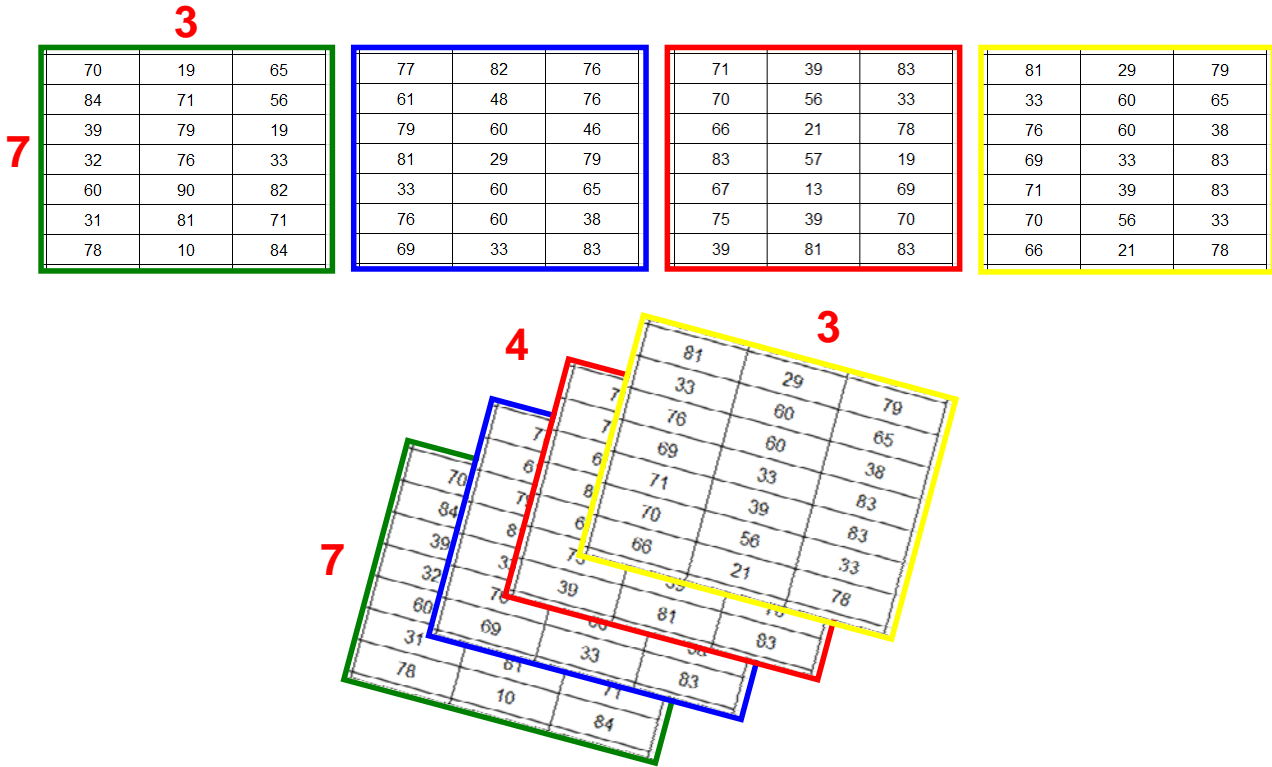
70	19	65
84	71	56
39	79	19
32	76	33
60	90	82
31	81	71
78	10	84
77	82	76
61	48	76
79	60	46
81	29	79
33	60	65
76	60	38
69	33	83
71	39	83
70	56	33

8

3

70	84	39	32	60	31	78	77
19	71	79	76	90	81	10	82
65	56	19	33	82	71	84	76

■ 三維以上的數據



■ 不同類型的資料

	Sepal.Length	Sepal.width	Petal.Length	Petal.width	Species
45	5.1	3.8	1.9	0.4	setosa
46	4.8	3.0	1.4	0.3	setosa
47	5.1	3.8	1.6	0.2	setosa
48	4.6	3.2	1.4	0.2	setosa
49	5.3	3.7	1.5	0.2	setosa
50	5.0	3.3	1.4	0.2	setosa
51	7.0	3.2	4.7	1.4	versicolor
52	6.4	3.2	4.5	1.5	versicolor
53	6.9	3.1	4.9	1.5	versicolor
54	5.5	2.3	4.0	1.3	versicolor
55	6.5	2.8	4.6	1.5	versicolor

	District	Group	Age	Holders	Claims
1	1	<17	<25	197	38
2	1	<17	25-29	264	35
3	1	<17	30-35	246	20
4	1	<17	>35	1680	156
5	1	1-1.57	<25	284	63
6	1	1-1.57	25-29	536	84
7	1	1-1.57	30-35	696	89
8	1	1-1.57	>35	3582	400
9	1	1.5-27	<25	133	19
10	1	1.5-27	25-29	286	52
11	1	1.5-27	30-35	355	74
12	1	1.5-27	>35	1640	233
13	1	>27	<25	24	4
14	1	>27	25-29	71	18
15	1	>27	30-35	99	19
16	1	>27	>35	452	77
17	2	<17	<25	85	22
18	2	<17	25-29	139	19
19	2	<17	30-35	151	22
20	2	<17	>35	931	87

- 陣列與矩陣
- 列表
- 資料框
- 因子

陣列與矩陣

建立矩陣與性質

計算機程式設計 - 2016F
Chap 05: 多維度資料
Feng-Li Lian @ NTU-EE

- `A <- matrix(c(1.2, -6.5, -3.4, 4.3, 5.6, -2.1), nrow = 2, ncol = 3)`
by column
- `B <- matrix(c(1.2, -3.4, 5.6, -6.5, 4.3, -2.1), nrow = 2, ncol = 3,`
`byrow = TRUE)`
by row
- `class(A)`
- `dim(A)`
- `attributes(A)`
- `nrow(A)`
- `dim(A)[1]`
- `ncol(A)`
- `dim(A)[2]`

- 7

矩陣與向量

計算機程式設計 - 2016F
Chap 05: 多維度資料
Feng-Li Lian @ NTU-EE

- `A <- matrix(c(1.2, -6.5, -3.4, 4.3, 5.6, -2.1), nrow = 2, ncol = 3)`
- `u <- as.numeric(A)`
- `v <- c(A)`
- `dim(v)`
- `length(v)`
- `nrow(A) * ncol(A)`
- `length(A)`

- 8

- `matrix(0, nrow = 2, ncol = 3)`
- `matrix(0, nrow = 3, ncol = 3)`
- `diag(0, nrow = 3)`
- `diag(3)`
- `diag(2.5, nrow = 3)`
- `diag(c(1, 2, 3), nrow = 3)`
- `diag(c(1, 2, 3))`

矩陣的加減乘

- `A <- matrix(1:9, nrow = 3, ncol = 3)`
- `B <- matrix(1:9, nrow = 3, ncol = 3, byrow = TRUE)`
- `A + B`
- `A - B`

- `A %*% B`

	[, 1]	[, 2]	[, 3]		[, 1]	[, 2]	[, 3]
[1,]	1	4	7	[1,]	1	2	3
[2,]	2	5	8	[2,]	4	5	6
[3,]	3	6	9	[3,]	7	8	9

- `A * B`

	[, 1]	[, 2]	[, 3]		[, 1]	[, 2]	[, 3]
[1,]	1	4	7	[1,]	1	2	3
[2,]	2	5	8	[2,]	4	5	6
[3,]	3	6	9	[3,]	7	8	9

用指標的方式取出某個分量

計算機程式設計 - 2016F
Chap 05: 多維度資料
Feng-Li Lian @ NTU-EE

- `A <- matrix(1:9, nrow = 3, ncol = 3)`
- `B <- matrix(1:9, nrow = 3, ncol = 3, byrow = TRUE)`
- `A[2, 3]`
- `A[1, 2]`
- `A[4]`
- `c(A)[4]`

- `R1 <- A[1,]`
- `class(R1)`

- `R1[1, 2]`

`[1]`

1	4	7
---	---	---

	[, 1]	[, 2]	[, 3]
[1,]	1	2	3
[2,]	4	5	6
[3,]	7	8	9

- `R1 * B`

`[1]`

1	4	7
---	---	---

	[, 1]	[, 2]	[, 3]
[1,]	1	2	3
[2,]	4	5	6
[3,]	7	8	9

- `R1 %*% B`

- 11

成為一個矩陣

計算機程式設計 - 2016F
Chap 05: 多維度資料
Feng-Li Lian @ NTU-EE

- `A <- matrix(1:9, nrow = 3, ncol = 3)`
- `B <- matrix(1:9, nrow = 3, ncol = 3, byrow = TRUE)`

- `RR1 <- A[1, , drop = FALSE]`

- `RR1[1, 2]`

- `RR1 %*% B`

- `as.matrix(R1)`

- 12

形成另外一個矩陣或向量

計算機程式設計 - 2016F
Chap 05: 多維度資料
Feng-Li Lian @ NTU-EE

- `A <- matrix(1:9, nrow = 3, ncol = 3)`
- `E <- A[c(1, 3),]`
- `class(E)`
- `F <- A[c(1, 3), 2]`
- `class(F)`

- 13

結合矩陣

計算機程式設計 - 2016F
Chap 05: 多維度資料
Feng-Li Lian @ NTU-EE

- `C <- matrix(1:4, nrow = 2, ncol = 2)`
- `D <- matrix(1:6, nrow = 2, ncol = 3)`
- `cbind(C, D)`
- `E <- matrix(1:4, nrow = 2, ncol = 2)`
- `F <- matrix(1:6, nrow = 3, ncol = 2)`
- `rbind(E, F)`

- 14

- `A <- matrix(1:9, nrow = 3, ncol = 3)`
- `t(A)`
- `t(A) %*% A`
- `diag(A)`
- `sum(diag(A))`

- `A <- matrix(c(1, 0, 0, 3, 0.5, 0, 2, 1, 0.25), nrow = 3, ncol = 3)`
- `det(A)`
- `Ainv <- solve(A)` A^{-1}
- `Ainv %*% A`

線性方程式： $Ax = b$

計算機程式設計 - 2016F
Chap 05: 多維度資料
Feng-Li Lian @ NTU-EE

- `A <- matrix(c(1, 0, 0, 3, 0.5, 0, 2, 1, 0.25), nrow = 3, ncol = 3)`
- `b <- c(2, 1, 3)`
- `solve(A, b)`

$$Ax = b$$

$$x = A^{-1} b$$

- 17

大綱

計算機程式設計 - 2016F
Chap 05: 多維度資料
Feng-Li Lian @ NTU-EE

列表

18


- `camera <- list(c("Leica", "Pentax", "Olympus", "Nikon"), c(1.2, 3.4), c("red", "green", "blue"))`
- `camera <- list(c("Leica", "Pentax", "Olympus", "Nikon"), c(1.2, 3.4), c("red", "green", "blue"))`
- `camera <- list(brand = c("Leica", "Pentax", "Olympus", "Nikon"), real.number = c(1.2, 3.4), color = c("red", "green", "blue"))`
- `camera <- list(brand = c("Leica", "Pentax", "Olympus", "Nikon"), real.number = c(1.2, 3.4), color = c("red", "green", "blue"))`

- 19

索引方式 1 - 使用 []

- `a1 <- camera[1]`
- `a1`
- `camera["brand"]`
- `class(a1)`

- 20



索引方式 2 – 使用 [[]]

計算機程式設計 – 2016F
Chap 05: 多維度資料
Feng-Li Lian @ NTU-EE

- `a2 <- camera[[1]]`
- `a2`

- `camera[["brand"]]`

- `class(a2)`

- 21



索引方式 3 – 使用 \$

計算機程式設計 – 2016F
Chap 05: 多維度資料
Feng-Li Lian @ NTU-EE

- `a3 <- camera$brand`
- `a3`

- `class(a3)`

- 22

- a1 # a1 是列表
- a2 # a2 是向量
- a3 # a3 是向量

- class(a1)
- class(a2)
- class(a3)

- camera[1][1]
- a1[1]

- a2[c(1, 2)]

- a3[2]

資料框

資料框 (data frame)

計算機程式設計 - 2016F
Chap 05: 多維度資料
Feng-Li Lian @ NTU-EE

- `x1 <- c("father", "mother", "brother", "sister")`
- `x2 <- c("Leica", "Pentax", "Olympus", "Nikon")`
- `x3 <- c("gold", "red", "green", "blue")`
- `x4 <- c(2, 1, 1, 2)`
- `camera <- data.frame(member = x1, brand = x2, color = x3, amount = x4)`
- `camera`

- 25

資料框 - 特徵

計算機程式設計 - 2016F
Chap 05: 多維度資料
Feng-Li Lian @ NTU-EE

- `camera`
- `class(camera)`
- `names(camera)`
- `colnames(camera)` # column names
- `rownames(camera)` # row names

- 26

- camera
- camera\$brand
- camera[, 2]
- camera[, "brand"]

- x5 <- c(8, 3, 2, 2)
- camera\$cost <- x5
- camera

- `test <- camera`
- `colnames(test)[c(4, 5)] <- c("number", "money")`
- `test`

- # 品牌 (brand) 為 Leica 的資料
- `camera[camera$brand == "Leica",]`
- `subset(camera, brand == "Leica")`
- # 品牌 (brand) 為 Leica 或 Nikon 的資料
- `camera[camera$brand %in% c("Leica", "Nikon"),]`
- `subset(camera, brand %in% c("Leica", "Nikon"))`
- # 價格 (cost) 大於 2 的資料
- `camera[camera$cost > 2,]`
- `subset(camera, cost > 2)`

- `A <- matrix(c(1.2, -6.5, -3.4, 4.3, 5.6, -2.1), nrow = 2, ncol = 3)`
- `rownames(A)` # row names
- `colnames(A)` # column names
- `D <- as.data.frame(A)`
- `names(D)`
- `colnames(D)`
- `rownames(D)`
- `D$V1`

因子

- 因子是用來描述具有等級與類別屬性的資料，例如：性別，階級等
- `x <- c("R", "G", "B", "R", "R", "B", "R", "G", "G")`
- `class(x)`
- `y <- factor(x)`
- `class(y)`

- # 這些等級可以用整數來表示
- `as.integer(y)`
- `levels(y)`
- `levels(y)[2]`
- `nlevels(y)`
- `levels(y)[as.integer(y)]`

因子 - 5的等級，3個分量

計算機程式設計 - 2016F
Chap 05: 多維度資料
Feng-Li Lian @ NTU-EE

- `gl(5, 3)` # factor levels up to 5 with repeats of 3
- `gl(n = 5, k = 3)`
- `class(gl(5, 3))`

- `gl(5, 2, 13)`
- `gl(n = 5, k = 2, length = 13)`
- `is.factor(gl(5, 2, 13))`

- 35

大綱

計算機程式設計 - 2016F
Chap 05: 多維度資料
Feng-Li Lian @ NTU-EE

作業

HW04：多維度資料格式

On 10/18, 2016

- 假設您預計購買一個手機，
從網路上找到一些手機的資料，
如下表所示：

廠牌	Brand	Apple	HTC	HTC	ASUS	ASUS
型號	Model	iPhone7	OneM8	OneS9	ZenFoneDeluxe	ZenFoneZoom
價格 (元)	Price	24500	21900	9990	8990	15990
螢幕 (吋)	Screen	4.7	5	5	5.5	5.5
重量 (克)	Weight	138	160	158	170	185
記憶體 (GB)	GB	32	16	32	16	64
日期 (年月日)	Date	20160916	20140328	20160617	20150827	20151201

- 37

HW04：多維度資料格式

On 10/18, 2016

- 編輯一個程式於 .R 檔，完成下面的工作：
 - 建立一個數列：Brand，放置五個手機的廠牌資料
 - 建立一個數列：Model，放置五個手機的型號資料
 - 建立一個數列：Price，放置五個手機的價格資料
 - 建立一個數列：Screen，放置五個手機的螢幕資料
 - 建立一個數列：Weight，放置五個手機的重量資料
 - 建立一個數列：GB，放置五個手機的記憶體資料
 - 建立一個數列：Date，放置五個手機的日期資料
 - 建立一個 5x3 的矩陣 (matrix)：Number，放置五個手機的價格，螢幕，重量三種資料
 - 建立一個 資料框 (data.frame)：Phone，放置這五個手機的七種資料
 - 建立一個 資料框 (data.frame)：PhoneCheap，放置這五個手機，其價格小於 10000 元的手機的所有資料
- 您可以從一個一個數列慢慢建立起，
也可以先建立一個資料框，再指定出個別的數列或矩陣
- 把執行的過程，以及產生的數據等，整理到報告檔 (pptx or pdf)。

- 38

HW04：多維度資料格式

On 10/18, 2016

- 繳交下面檔案，檔案名稱：[HW04_學號](#)
 - 主要指定檔案：[HW04_B01921001.R](#)
 - 報告檔案：[HW04_B01921001.pptx](#) 或者 [.pdf](#)
- 繳交方式與期限：
 - E-mail 上面四個檔案到：ntucp105f@gmail.com
 - E-mail 主旨：[HW04_B01921001](#) (就是，作業編號_您的學號)
 - 繳交期限：**10/19 (Wed), 2016, 11pm 以前**
 - **HW04**，每位繳交自己的作業，或者兩人繳交一份
 - 兩人完成作業者，請註明工作模式，例如：
 - A 進行指令測試與報告編輯，然後給 B 重複測試
 - A 進行指令測試，B 進行報告編輯
 - A 進行指令測試與報告編輯，B 僅看過報告，但沒有實際測試
 - A 進行指令測試與報告編輯，B 沒有參與指令測試，也沒有看過報告