

2017 Big Data – R programming Homework 2

2017/09/30

范真璋

Part I: Vector & Matrix

1. 將下列數字新增成一向量 V

35, 4, 12, 2, 14, 11, 9, 5, 11, 3, 38, 12, 1, 0, 4, 2

```
> V <- c(35, 4, 12, 2, 14, 11, 9, 5, 11, 3, 38, 12, 1, 0, 4, 2)
> V
[1] 35  4 12  2 14 11  9  5 11  3 38 12  1  0  4  2
```

2. 將向量 V 新增成一 4 X 4 的矩陣 A，以 by column 排列

```
> A <- matrix(V, nrow = 4, ncol = 4)
> A
      [,1] [,2] [,3] [,4]
[1,]   35   14   11    1
[2,]    4   11    3    0
[3,]   12    9   38    4
[4,]    2    5   12    2
```

3. 求下列矩陣及其行列式

a. A^{-1}

```
> Aa <- solve(A)
> Aa
      [,1]      [,2]      [,3]      [,4]
[1,] 0.034811530 -0.03680710 -0.004545455 -0.008314856
[2,] -0.007095344  0.09667406 -0.018181818  0.039911308
[3,] -0.020399113  0.02793792  0.072727273 -0.135254989
[4,]  0.105321508 -0.37250554 -0.386363636  1.220066519
> det(Aa)
[1] 0.0001108647
```

b. AA^T

```
> Ab <- A %*% t(A)
> Ab
      [,1] [,2] [,3] [,4]
[1,] 1543  327  968  274
[2,]  327  146  261   99
[3,]  968  261 1685  533
[4,]  274   99  533  177
> det(Ab)
[1] 81360400
```

c. $A^T A$

```
> Ac <- t(A) %*% A
> Ac
      [,1] [,2] [,3] [,4]
[1,] 1389  652  877   87
[2,]  652  423  589   60
[3,]  877  589 1718  187
[4,]   87   60  187   21
> det(Ac)
[1] 81360400
```

Part II: DataFram & List

1. 使用 read_excel()讀取 AQI.xlsx，並將資料放入 aqi 變數中

```
> library("readxl")
> aqi <- read_excel("D:/三/上/Big Data/hw/hw2/AQI.xlsx")
> aqi
# A tibble: 76 x 20
  SiteName County AQI Pollutant Status SO2 CO CO_8hr O3 O3_8hr PM10 PM2.5 NO2 NOX NO
  <chr> <chr> <dbl> <chr> <chr> <dbl> <dbl> <dbl> <chr> <dbl> <dbl> <chr> <dbl> <dbl> <dbl>
1 二林 彰化縣 54 臭氧八小時 普通 3.8 0.26 0.3 57 56 47 14 9.2 9.7 0.5
2 三重 新北市 43 <NA> 良好 2.6 1.56 1.2 - NA 30 4 46.0 98.0 52.0
3 三義 苗栗縣 58 細懸浮微粒 普通 1.5 0.32 0.3 40 36 48 31 NA NA NA
4 土城 新北市 38 <NA> 良好 5.3 0.45 0.4 35 41 13 9 20.0 22.0 1.9
5 士林 臺北市 40 <NA> 良好 1.0 0.33 0.3 39 43 13 8 14.0 16.0 1.6
6 大同 臺北市 48 <NA> 良好 3.5 1.87 1.0 - NA 32 3 51.0 129.0 78.0
7 大里 臺中市 76 細懸浮微粒 普通 2.6 0.49 0.5 36 47 67 31 28.0 30.0 2.7
8 大園 桃園市 46 <NA> 良好 2.7 0.24 0.2 46 50 36 9 16.0 18.0 2.8
9 大寮 高雄市 100 細懸浮微粒 普通 4.2 0.33 0.6 53 57 54 23 24.0 26.0 2.2
10 小港 高雄市 102 細懸浮微粒 對敏感族群不良 4.7 0.30 0.4 57 57 54 20 24.0 26.0 1.5
# ... with 66 more rows, and 5 more variables: windSpeed <dbl>, windDirec <dbl>, PublishTime <dtm>, PM2.5_AVG <dbl>,
# PM10_AVG <dbl>
```

2. 利用 naming 做 selection，從 aqi 中選出

"SiteName", "County", "AQI", "Status", "CO", "O3", "PM2.5"等資訊

```
> new_aqi <- aqi[c("SiteName", "County", "AQI", "Status", "CO", "O3", "PM2.5")]
> new_aqi
# A tibble: 76 x 7
   SiteName County    AQI    Status    CO    O3 PM2.5
   <chr>    <chr>    <dbl>    <chr>    <dbl> <chr> <chr>
1 二林 彰化縣    54    普通    0.26    57    14
2 三重 新北市    43    良好    1.56     -     4
3 三義 苗栗縣    58    普通    0.32    40    31
4 土城 新北市    38    良好    0.45    35     9
5 士林 臺北市    40    良好    0.33    39     8
6 大同 臺北市    48    良好    1.87     -     3
7 大里 臺中市    76    普通    0.49    36    31
8 大園 桃園市    46    良好    0.24    46     9
9 大寮 高雄市   100    普通    0.33    53    23
10 小港 高雄市   102 對敏感族群不良 0.30    57    20
# ... with 66 more rows
```

3. 製作 list，以每 10 行為一單位，將 aqi 切割成 7 組 datafram 並給

予名稱 first, second, third……存入 list 中，最後一組名稱為 last

```
> aqi_df1 <- new_aqi[1:10, 1:7]
> aqi_df2 <- new_aqi[11:20, 1:7]
> aqi_df3 <- new_aqi[21:30, 1:7]
> aqi_df4 <- new_aqi[31:40, 1:7]
> aqi_df5 <- new_aqi[41:50, 1:7]
> aqi_df6 <- new_aqi[51:60, 1:7]
> aqi_df7 <- new_aqi[61:70, 1:7]
> aqi_df8 <- new_aqi[71:76, 1:7]
> aqi_list <- list("first"=aqi_df1, "second"=aqi_df2, "third"=aqi_df3, "fourth"=aqi_df4, "fifth"=aqi_df5,
+               "sixth"=aqi_df6, "seventh"=aqi_df7, "last"=aqi_df8)
> aqi_list
```

```
$first
# A tibble: 10 x 7
  SiteName County   AQI      Status    CO    O3 PM2.5
  <chr>   <chr> <dbl>    <chr> <dbl> <chr> <chr>
1 二林 彰化縣    54      普通  0.26   57   14
2 三重 新北市    43      良好  1.56    -    4
3 三義 苗栗縣    58      普通  0.32   40   31
4 土城 新北市    38      良好  0.45   35    9
5 士林 臺北市    40      良好  0.33   39    8
6 大同 臺北市    48      良好  1.87    -    3
7 大里 臺中市    76      普通  0.49   36   31
8 大園 桃園市    46      良好  0.24   46    9
9 大寮 高雄市   100      普通  0.33   53   23
10 小港 高雄市   102 對敏感族群不良  0.30   57   20
```

```
$second
# A tibble: 10 x 7
  SiteName County   AQI      Status    CO    O3 PM2.5
  <chr>   <chr> <dbl>    <chr> <dbl> <chr> <chr>
1 中山 臺北市    31      良好  0.60   23 <NA>
2 中壢 桃園市    55      普通  1.34   12   21
3 仁武 高雄市   133 對敏感族群不良  0.25   60   33
4 斗六 雲林縣   131 對敏感族群不良  0.43   51   40
5 冬山 宜蘭縣    39      良好  0.22   45    6
6 古亭 臺北市    36      良好  0.46   25   10
7 左營 高雄市   118 對敏感族群不良  0.25   68   21
8 平鎮 桃園市    35      良好  0.48   27    8
9 永和 新北市    33      良好  0.59   24    9
10 安南 臺南市    77      普通  0.28   67   17
```

```
$third
# A tibble: 10 x 7
  SiteName County   AQI      Status    CO    O3 PM2.5
  <chr>   <chr> <dbl>    <chr> <dbl> <chr> <chr>
1 朴子 嘉義縣    65      普通  0.23   63   13
2 汐止 新北市    53      普通  0.50   11   19
3 竹山 南投縣   144 對敏感族群不良  0.44   48   50
4 竹東 新竹縣    56      普通  0.53   14   21
5 西屯 臺中市    65      普通  0.32   43   24
6 沙鹿 臺中市    57      普通  0.28   48   16
7 宜蘭 宜蘭縣    33      良好  0.32   34    3
8 忠明 臺中市    77      普通  0.41   38   30
9 松山 臺北市    40      良好  0.50   14   16
10 板橋 新北市    33      良好  0.53   22   12
```

\$fourth

A tibble: 10 x 7

	SiteName	County	AQI	Status	CO	O3	PM2.5
	<chr>	<chr>	<dbl>	<chr>	<dbl>	<chr>	<chr>
1	林口	新北市	49	良好	0.49	25	14
2	林園	高雄市	108	對敏感族群不良	0.23	74	12
3	花蓮	花蓮縣	47	良好	0.43	44	7
4	金門	金門縣	83	普通	0.29	47	28
5	前金	高雄市	90	普通	0.34	66	13
6	前鎮	高雄市	111	對敏感族群不良	0.31	58	23
7	南投	南投縣	132	對敏感族群不良	0.53	34	40
8	屏東	屏東縣	153	對所有族群不良	0.47	57	43
9	恆春	屏東縣	49	良好	0.15	52	13
10	美濃	高雄市	121	對敏感族群不良	0.45	75	43

\$fifth

A tibble: 10 x 7

	SiteName	County	AQI	Status	CO	O3	PM2.5
	<chr>	<chr>	<dbl>	<chr>	<dbl>	<chr>	<chr>
1	苗栗	苗栗縣	66	普通	0.37	38	25
2	埔里	南投縣	152	對所有族群不良	0.47	60	65
3	桃園	桃園市	36	良好	0.42	31	9
4	馬公	澎湖縣	44	良好	0.22	51	20
5	馬祖	連江縣	56	普通	0.23	48	18
6	基隆	基隆市	38	良好	0.32	44	9
7	崙背	雲林縣	80	普通	0.23	62	29
8	淡水	新北市	35	良好	0.38	34	11
9	麥寮	雲林縣	85	普通	0.23	56	28
10	善化	臺南市	104	對敏感族群不良	0.26	61	32

\$sixth

A tibble: 10 x 7

	SiteName	County	AQI	Status	CO	O3	PM2.5
	<chr>	<chr>	<dbl>	<chr>	<dbl>	<chr>	<chr>
1	復興	高雄市	89	普通	0.54	61	16
2	湖口	新竹縣	46	良好	0.25	44	15
3	萊寮	新北市	38	良好	0.55	32	9
4	陽明	臺北市	48	良好	0.17	52	ND
5	新竹	新竹市	51	普通	0.42	29	17
6	新店	新北市	39	良好	0.45	35	7
7	新莊	新北市	36	良好	0.51	27	4
8	新港	嘉義縣	84	普通	0.23	61	23
9	新營	臺南市	96	普通	0.30	57	26
10	楠梓	高雄市	118	對敏感族群不良	0.34	62	33

\$seventh

A tibble: 10 x 7

	SiteName	County	AQI	Status	CO	O3	PM2.5
	<chr>	<chr>	<dbl>	<chr>	<dbl>	<chr>	<chr>
1	萬里	新北市	46	良好	0.21	56	15
2	萬華	臺北市	33	良好	0.55	20	5
3	嘉義	嘉義市	113	對敏感族群不良	0.37	58	34
4	彰化	彰化縣	72	普通	0.41	40	23
5	臺西	雲林縣	60	普通	0.19	56	26
6	臺東	臺東縣	30	良好	0.43	31	6
7	臺南	臺南市	80	普通	0.36	60	26
8	鳳山	高雄市	115	對敏感族群不良	0.54	51	26
9	潮州	屏東縣	123	對敏感族群不良	0.42	70	7
10	線西	彰化縣	73	普通	0.36	52	29

\$last

A tibble: 6 x 7

	SiteName	County	AQI	Status	CO	O3	PM2.5
	<chr>	<chr>	<dbl>	<chr>	<dbl>	<chr>	<chr>
1	橋頭	高雄市	114	對敏感族群不良	0.31	48	26
2	頭份	苗栗縣	55	普通	0.42	33	17
3	龍潭	桃園市	43	良好	0.49	23	15
4	豐原	臺中市	57	普通	0.38	43	19
5	關山	臺東縣	31	良好	NA	35	9
6	觀音	桃園市	47	良好	0.23	46	8