

## HW #1

ST 437 / 537

① Let  $\underline{x} = \begin{pmatrix} 5 \\ 1 \\ 3 \end{pmatrix}$  and  $\underline{y} = \begin{pmatrix} -1 \\ 3 \\ 1 \end{pmatrix}$

437/537

a. Do 'by hand' the following:

1. Find the lengths of  $\underline{x}$  and  $\underline{y}$

2. Calculate  $\underline{x}^T \underline{y}$

3. Are  $\underline{x}$  and  $\underline{y}$  orthogonal? Explain your answer

b. Repeat the previous questions in R. Provide code

437/537

② Let  $X$  denote a  $n \times n$  square matrix and  $\underline{y} \in \mathbb{R}^n$  an  $n$ -tuple ~~matrix~~ vector of all ones,  $\underline{y} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$ .

a. Interpret  $\underline{y}^T X$ .

b. Interpret  $X \underline{y}$

437/537

③ Consider the matrix  $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$

a. Find the trace, determinant and inverse of this matrix

b. Multiply this matrix by its inverse and verify that your solution is correct

Work a. and b first by hand and then in R.  
Show your work and provide code.

c. Find the eigenvalues of  $A$ . Provide R code.

537

(4) Let  $Z$  denote an arbitrary  $n \times m$  matrix

a. Show that  $Z^T Z$  is symmetric

b. Show that  $Z^T Z$  is nonnegative ~~definite~~ definite matrix

537

(5) Suppose that  $X$  and  $Y$  are independent normal random variables, with the same variance.

Show that  $X+Y$  and  $X-Y$  are also independent.

137/537

(6) Consider the data in the Table 1.5 of the attached pdf

a. Construct side-by-side boxplots of all the variables

b. Compute summary statistics (using the function `summary(.)` in R) for each variable

c. Construct a "pairs-plot"

d. Denote by  $\tilde{x}_i = \begin{pmatrix} x_{i1} \\ \vdots \\ x_{i7} \end{pmatrix}$  each row-entry in the table

Compute  $\bar{X}$  and  $S$ , the mean <sup>vector</sup> and the  $7 \times 7$  Covariance matrix.

Provide R code for each part clearly.

1.6. The data in Table 1.5 are 42 measurements on air-pollution variables recorded at 12:00 noon in the Los Angeles area on different days. (See also the air-pollution data on the web at [www.prenhall.com/statistics](http://www.prenhall.com/statistics).)

- (a) Plot the marginal dot diagrams for all the variables.
- (b) Construct the  $\bar{x}$ ,  $S_n$ , and  $R$  arrays, and interpret the entries in  $R$ .

**Table 1.5** Air-Pollution Data

Wind ( $x_1$ )	Solar radiation ( $x_2$ )	CO ( $x_3$ )	NO ( $x_4$ )	NO <sub>2</sub> ( $x_5$ )	O <sub>3</sub> ( $x_6$ )	HC ( $x_7$ )
8	98	7	2	12	8	2
7	107	4	3	9	5	3
7	103	4	3	5	6	3
10	88	5	2	8	15	4
6	91	4	2	8	10	3
8	90	5	2	12	12	4
9	84	7	4	12	15	5
5	72	6	4	21	14	4
7	82	5	1	11	11	3
8	64	5	2	13	9	4
6	71	5	4	10	3	3
6	91	4	2	12	7	3
7	72	7	4	18	10	3
10	70	4	2	11	7	3
10	72	4	1	8	10	3
9	77	4	1	9	10	3
8	76	4	1	7	7	3
8	71	5	3	16	4	4
9	67	4	2	13	2	3
9	69	3	3	9	5	3
10	62	5	3	14	4	4
9	88	4	2	7	6	3
8	80	4	2	13	11	4
5	30	3	3	5	2	3
6	83	5	1	10	23	4
8	84	3	2	7	6	3
6	78	4	2	11	11	3
8	79	2	1	7	10	3
6	62	4	3	9	8	3
10	37	3	1	7	2	3
8	71	4	1	10	7	3
7	52	4	1	12	8	4
5	48	6	5	8	4	3
6	75	4	1	10	24	3
10	35	4	1	6	9	2
8	85	4	1	9	10	2
5	86	3	1	6	12	2
5	86	7	2	13	18	2
7	79	7	4	9	25	3
7	79	5	2	8	6	2
6	68	6	2	11	14	3
8	40	4	3	6	5	2

Source: Data courtesy of Professor G. C. Tiao.

mpact with age  
rs:

11  
3.99

$s_{11}$  and  $s_{22}$ . Com-  
ment  $r_{12}$ . Interpret

array  $S_n$ , and the

$x_3$  = assets  
(billions)

1,484.10  
750.33  
766.42  
1,110.46  
1,031.29  
195.26  
193.83  
191.11  
1,175.16  
211.15

2000,

ables  $x_1$  and  $x_2$ . Com-

$(x_1, x_3)$ . Comment on