EP375 Computational Physics

Topic 2

ARRAY MANIPULATION & DATA ANALYSIS

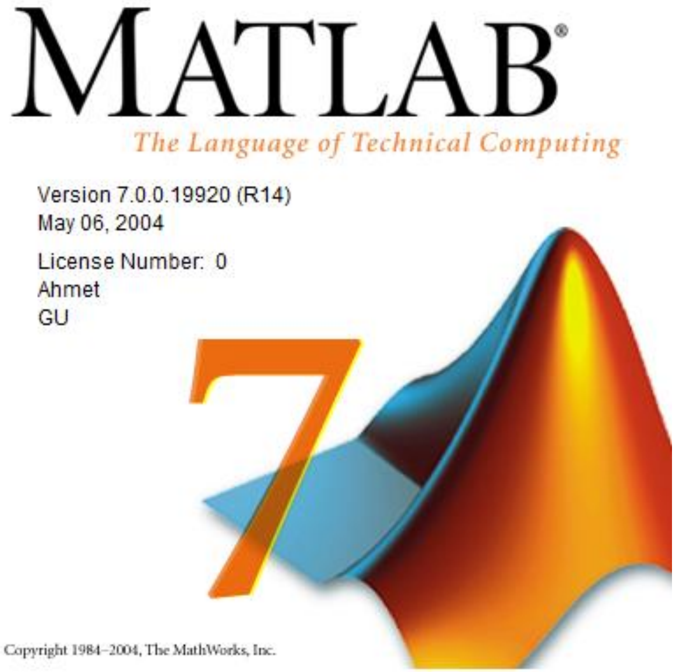
Department of Engineering Physics

University of Gaziantep

Feb 2016

Content

1. **1D Arrays**
2. **2D Arrays**
3. **Array Functions**
4. **Reading Text files**
5. **Data Analysis**



OneDim Arrays (Vectors)

- We learned before that an array can be created in may ways

```
>> x = [0  0.25  0.5  0.75  1]
x = 0    0.2500    0.5000    0.7500    1.0000
```

```
>> x = 0:0.25:1
x = 0    0.2500    0.5000    0.7500    1.0000
```

```
>> for i=1:5
     x(i) = (i-1)*0.25;
   end
>> x
x = 0    0.2500    0.5000    0.7500    1.0000
```

```
>> dizi = 1:7
dizi = 1      2      3      4      5      6      7
```

```
>> dizi2 = -5:2:5
dizi2 = -5     -3      -1      1      3       5
```

```
>> v = [1 2 3]    % row vector
v = 1   2   3
```

```
>> v = [1; 2; 3]   % column vector
v =
     1
     2
     3
```

```
>> v = [1 2 3]'   % transpose of a row vector
v =
     1
     2
     3
```

TwoDim Arrays (Matrices)

```
>> A = [1 1 1; 2 2 2]

A =   1      1      1
      2      2      2
```

```
>> A = [1 1 1
        2 2 2]

A =   1      1      1
      2      2      2
```

```
>> B = A'

B =   1      2
      1      2
      1      2
```

```
>> M = [9 8 7;  6 5 4;  3 2 1]
M =
     9     8     7
     6     5     4
     3     2     1
>> M(2,3)
ans = 4

>> M(:,3)
ans =
      7
      4
      1

>> M(2,:)
ans = 6     5     4

>> M(2,3) = 100
ans =
     9     8     7
     6     5   100
     3     2     1
```

Some Array Functions

`length(x)`	Number of elements of (vector) x
`sum(x)`	Sum the elements of x
`mean(x)`	Aritmetic mean of x
`std(x)`	Standard deviation of x
`prod(x)`	Product of the elements of x
`dot(x,y)`	Scalar product of x and y
`cross(x,y)`	Vector product of x and y
`linspace()`	Linearly spaced data
`logspace()`	Logaritmicly spaced data
`size(A)`	Number of rows and of matrix A
`zeros(n)`	nxn matrix whose elements are zero
`ones(n)`	nxn matrix whose elements are one
`eye(n)`	nxn unit matrix
`rand(n)`	nxn random matrix
`magic(n)`	nxn magic square

- **`linspace()`** function

```
>> x = 0:0.25:1
x = 0      0.2500     0.5000     0.7500     1.0000
```

```
>> x = linspace(0,1,5)
x = 0      0.2500     0.5000     0.7500     1.0000
```

- **`logspace()`** function: *logarithmic counterpart of linspace*

```
>> x = logspace(0,1,5)
x =  1.0000     1.7783     3.1623     5.6234    10.0000
```

creates 5 logarithmically spaced elements

starting with $x = 10^0$ and ending with $x = 10^1$.

```
>> x = [0  0.25  0.5  0.75 1];
>> length(x)
ans = 5
```

```
>> A = [1 1 1; 2 2 2];
>> size(A)
ans =  2   3
```

```
>> x = [1  2  2.5  3   3.1];
>> sum(x)
ans = 11.6000

>> prod(x)
46.5000
```

```
>> a = [1 2 4];
>> b = [0 2 5];
>> dot(a,b)
ans = 24
>> cross(a,b)
ans =  2    -5     2
```

zeros(m,n) *returns a matrix of m rows and n columns that is filled with zeroes*

ones(m,n) *returns a matrix of m rows and n columns that is filled with ones*

rand(m,n) *returns a matrix of m rows and n columns that is filled with uniform random number between [0,1]*

eye(n) *creates an n* x *n identity (unit) matrix.*

```
>> P = zeros(2,3)
P =  0      0      0
     0      0      0
>> P = ones(2,3)
P =  1      1      1
     1      1      1
>> P = rand(2,3)
P = 0.9501    0.6068    0.8913
    0.2311    0.4860    0.7621
>> I = eye(2)
I =  1      0
     0      1
```

Vectors and Relational Operators

- Recall the relational operators ==, ~=, <, >, <=, and >=.
- These operators can be applied to vectors element-wise to generate a same-sized vector of logical results.

```
>> X = [1 9 8 4];
>> Y = [4 3 0 4];
>> X>Y
ans = 0 1 1 0

>> X == Y
ans = 0 0 0 1

>> X>Y | Y == 4
ans = 1 1 1 1
```

```
>> X = [1 9 8 4];
>> Y = [4 3 0 4];
>> sum(X>Y)
ans = 2

>> sum(X == Y)
ans = 1

>> sum(X>Y | Y == 4)
ans = 4
```

Reading Text Files

Consider **gravity.txt** (*) contains 1000 measurements of gravitational acceleartion (g) on the Earth surface at sea level.

Find the mean, maximum and minimum values of the data.

We can read this data directly into a vector and process it as follows:

```
>> g = load ('gravity.txt');
>> mean(g)
ans = 9.8120
>> max(g)
ans = 10.4856
>> min(g)
ans = 9.0473
```

```
>> g = textread('gravity.txt');
>> mean(g)
ans = 9.8120
>> max(g)
ans = 10.4856
>> min(g)
ans = 9.0473
```

(*) Download the file at:

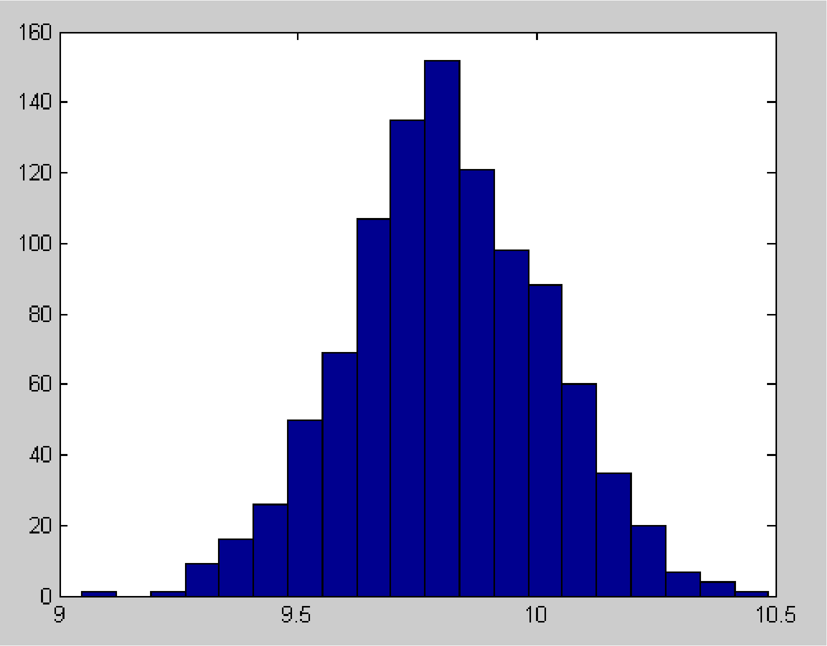
http://www1.gantep.edu.tr/~bingul/ep375/src/gravity.txt

Data Analysis

Data analysis is a very broad subject covering many techniques and types of data. In this lecture we will study some basic calculations that are commonly performed on sampled data.

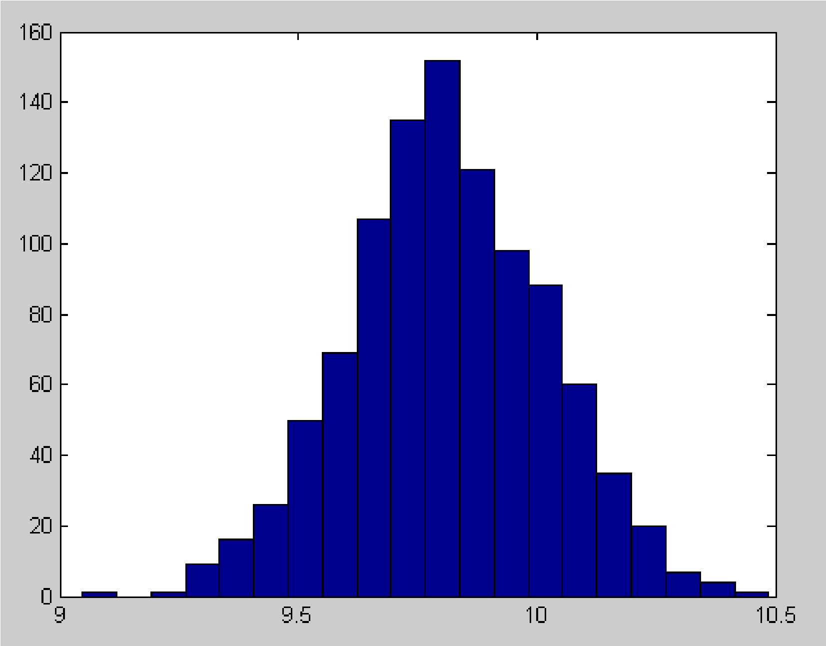
Consider again the file **gravity.txt**. We can plot the histogram of the data:

```
>> g = load('gravity.txt');
>> hist(g,20)
```



The ***mean*** $\overline{x}$ of the sample is given by:

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$



The ***standard deviation*** σ of the sample is defined as:

$$\sigma = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2}$$

```
>> g = load('gravity.txt');
>> mean(g)
ans = 9.8120
>> std(g)
ans = 0.2077
```

For this data, σ is the size of the variation of *g* about the mean.

For bi-variate data (two variables) the correlation coefficient (ρ) is a measure of the linear dependence between one variable and the other.

Given a sample (size n) of bi-variate data,
$Z = \{ (x_1,y_1), (x_2,y_2), (x_3,y_3), \ldots , (x_n,y_n) \}$

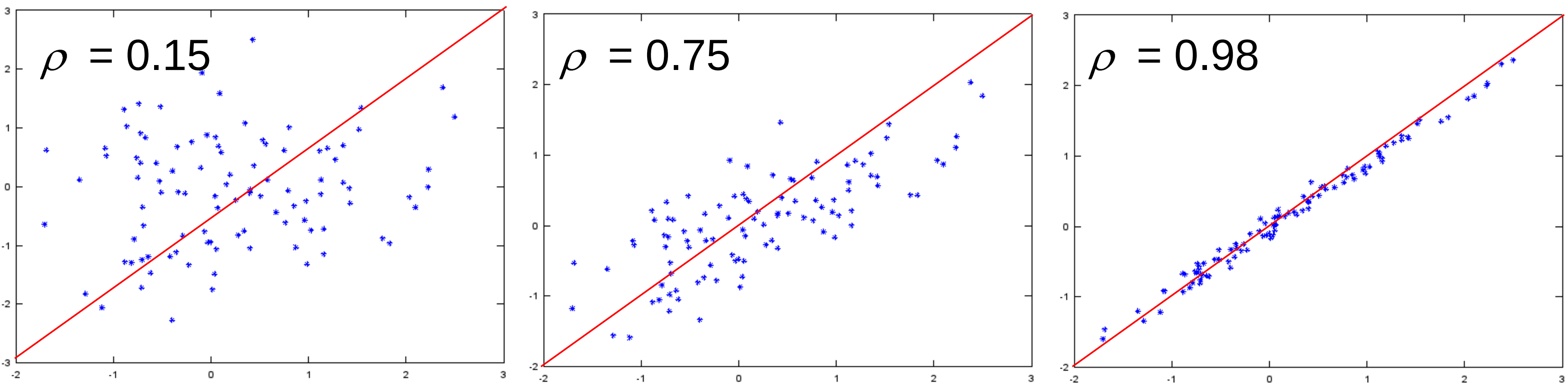
$$\boxed{\rho = \frac{\overline{xy} - \overline{x}.\overline{y}}{\sigma_x \sigma_y}}$$

$$\overline{xy} = \frac{1}{n}\sum_{i=1}^{n} x_i y_i \qquad \overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad \overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

$$\sigma_x = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2} \qquad \sigma_y = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(y_i - \overline{y})^2}$$

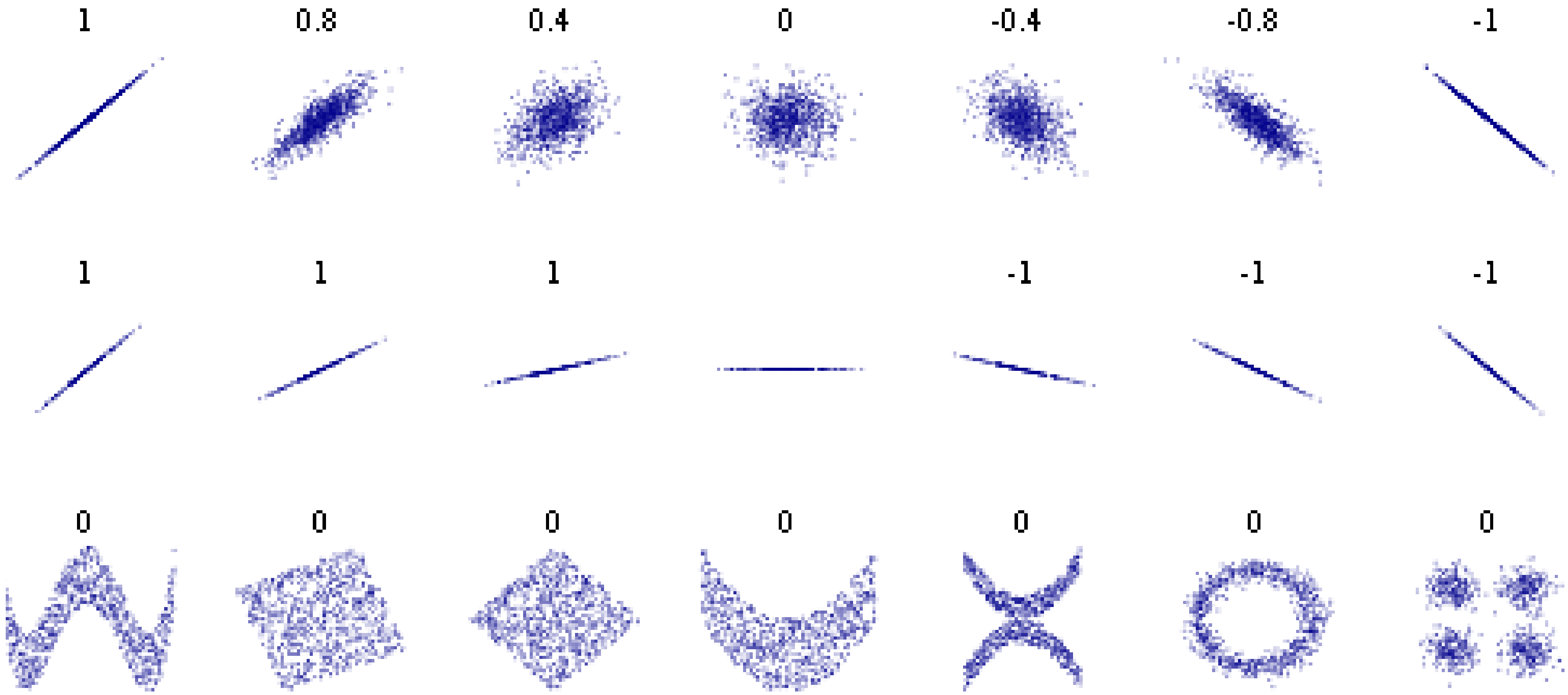
$$\boxed{\rho = \frac{\overline{xy} - \overline{x}.\overline{y}}{\sigma_x \sigma_y}}$$

$$-1 \leq \rho \leq 1$$



$\rho = 0$ if there is no correlation

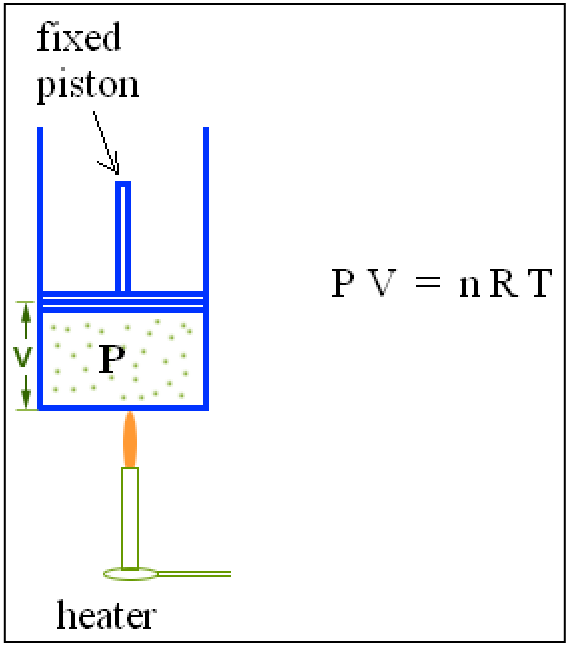
$\rho = \pm 1$ if X and Y are fully correlated


1
0.8
0.4
0
-0.4
-0.8
-1
1
1
1
-1
-1
-1
0
0
0
0
0
0
0

Consider a heated gas in a fixed volume.
The following data (stored in file TP.txt) contains
the temperature (T) in oC and
the pressure (P) in hPa
of the gas.

We can read this data directly
into a vector and process it as follows:

```
>> data = load('TP.txt');
>> T = data(:,1);
>> P = data(:,2);
>> plot(T,P,'*')
```



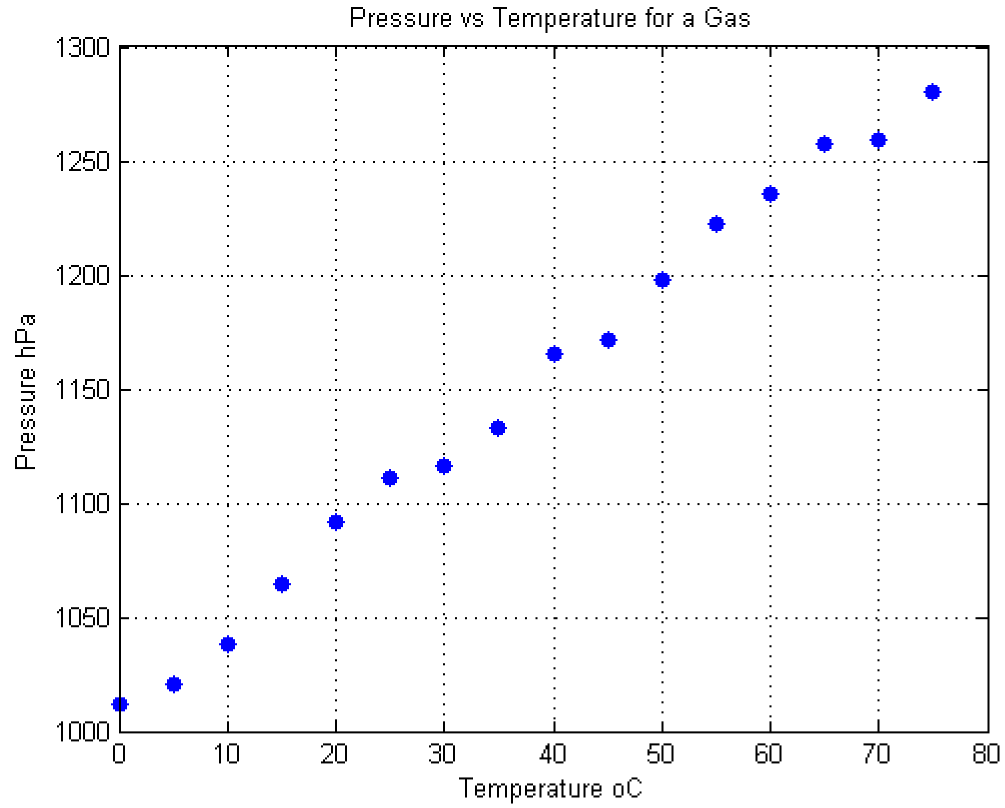
T(oC)	P(hPa)
0	1012
5	1021
10	1038
15	1065
20	1091
25	1111
30	1116
35	1133
40	1165
45	1171
50	1198
55	1223
60	1235
65	1258
70	1259
75	1280

Download the file at:

`http://www1.gantep.edu.tr/~bingul/ep375/src/TP.txt`

The correlation between

T and P:




```
>> data = load('TP.txt');
>> T = data(:,1);
>> P = data(:,2);
>> plot(T,P,'*')
>>
>>
>>
>>
>> rho = (mean(T.*P)-mean(T)*mean(P))/(std(T)*std(P))
rho = 0.94
```

There is clearly <u>strong</u> positive dependence on the temperature since $\rho \approx 1$. "A higher temperature leads to a higher pressure".

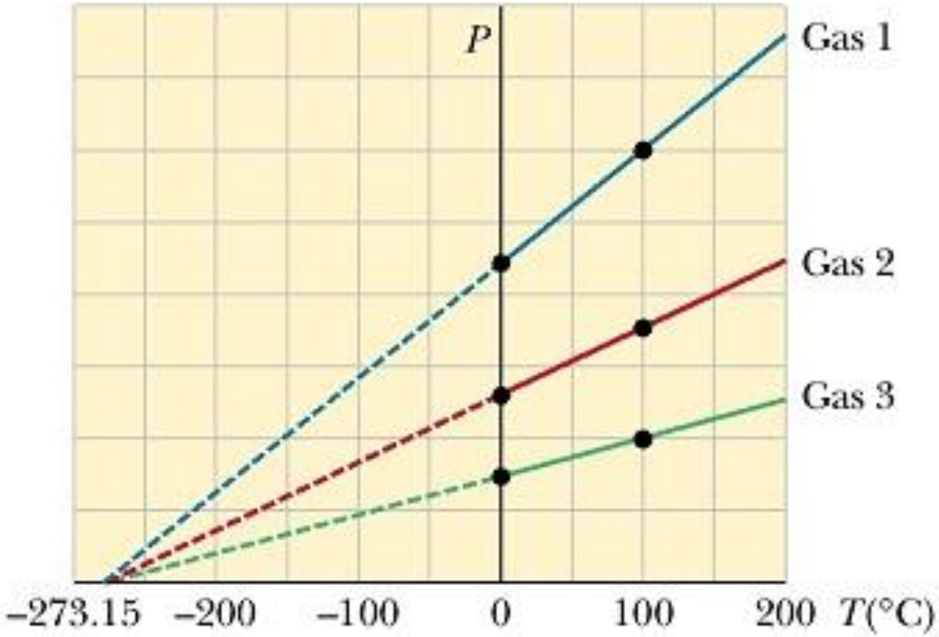


Figure 19.4 Pressure versus temperature for three dilute gases. Note that, for all gases, the pressure extrapolates to zero at the temperature -273.15°C.

Exercises

1. What is the output of the following program?

```
>> A = [1 2; 4 5];
>> det(A)
>> inv(A)
>> det(inv(A'))
```

2. What is the output of the following program?

```
>> a = [1 2 3];
>> b = [3 2 1];
>> log10(a)
>> a + b
>> 3*a
>> a.*b
>> a*b'
```

3.

Using the following data:

$$X = \{ 23, 31, 61, 17, 27, 51 \}$$
$$Y = \{ 16, 23, 49, 9, 14, 45 \}$$

Calculate

(a) the mean of X and Y

(b) the standard deviation of X and Y

(c) the correlation coefficient of X and Y

4.

An industrial refrigerator is used to cool food in a processsing factor. The factory engineer wishes to determine whether the temperature of the refrigerator is affected by the wind speed near it's external heat exchanger. An experiment is performed where the wind speed, W (m/s), and temperature, T(°C), of the refrigerator are sampled at 6 am in the morning every day over a period of twelve days.

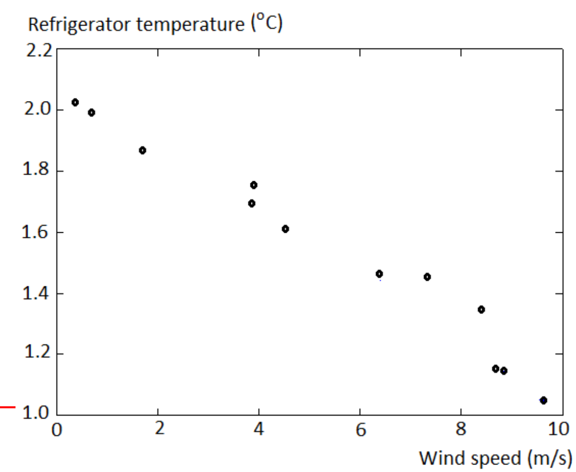
The results are given in the table . Use the table to calculate the correlation coefficient (comment on the results). Also try to plot Temperature versus WindSpeed to visualise the correlation.

W (m/s)	T(°C)
3.8	1.69
8.4	1.34
7.3	1.45
3.9	1.75
1.7	1.87
9.6	1.05
4.5	1.60
6.4	1.45
0.4	2.02
8.7	1.15
8.8	1.15
0.7	1.99

Answer: rho = −0.986

The results show that there is a strong negative correlation betweeen wind speed and refrigerator temperature.

"A higher the wind speed leads to a lower temperature".



5.

A local ice cream shop keeps track of how much ice cream they sell versus the temperature on that day. Table shows their data for the last 10 days. Calculate the correlation coefficient of the data and comment on the result..

Temperature	Ice Cream Sales (TL)
14 °C	215
16 °C	325
15 °C	332
18 °C	406
19 °C	412
25 °C	614
23 °C	544
18 °C	421
22 °C	445
17 °C	408

References

[1]. http://www.mathworks.com/products/matlab

[2]. Numerical Methods in Engineering with MATLAB,
J. Kiusalaas, Cambridge University Press (2005)

[3]. Numerical Methods for Engineers, 6th Ed.
S.C. Chapra, Mc Graw Hill (2010)

[4]. http://wikipedia.org/wiki/Mean

[5]. http://wikipedia.org/wiki/Standard_deviation

[6]. http://en.wikipedia.org/wiki/Correlation