

An Introduction to Matlab

Version 2.2

David F. Griffiths

Department of Mathematics
The University

Dundee DD1 4HN

With additional material by Ulf Carlsson Department of Vehicle Engineering KTH, Stockholm, Sweden

Contents			15 Examples in Plotting	
1	MATLAB	2	16 Matrices—Two–Dimensional Arrays	13
2	Starting Up 2.1 Windows Systems	2 2 2 2 3	16.1 Size of a matrix	13 14 14 14 14 15
3	Matlab as a Calculator	3	16.7 Tabulating Functions	15 15
4	Numbers & Formats	3	16.9 Dot product of matrices (.*) 16.10Matrix-vector products	16 16
5	Variables 5.1 Variable Names	3	16.11Matrix-Matrix Products	17 17
6	Suppressing output	3	17 Systems of Linear Equations 17.1 Overdetermined system of linear equa-	17
7	Built—In Functions 7.1 Trigonometric Functions	4 4	tions	18
	7.2 Other Elementary Functions	4	18 Characters, Strings and Text	19
8	Vectors 8.1 The Colon Notation	4 5	19 Loops	20
	8.2 Extracting Bits of a Vector	5	20 Logicals	21
	8.3 Column Vectors	5	20.1 While Loops	$\frac{22}{22}$
	8.4 Transposing	5		
9	Keeping a record	6	21 Function m-files 21.1 Examples of functions	23 24
10	Plotting Elementary Functions 10.1 Plotting—Titles & Labels 10.2 Grids 10.3 Line Styles & Colours 10.4 Multi-plots 10.5 Hold 10.6 Hard Copy	6 6 7 7 7 7 7 8	22 Further Built-in Functions 22.1 Rounding Numbers 22.2 The sum Function 22.3 max & min 22.4 Random Numbers 22.5 find for vectors 22.6 find for matrices	25 25 25 25 26 26
	10.8 Zooming	8	23 Plotting Surfaces	27
	10.10Controlling Axes	9	24 Timing	28
11	Keyboard Accelerators	9	25 On-line Documentation	28
12	Copying to and from Word and other applications 12.1 Window Systems	9 10 10	26 Reading and Writing Data Files 26.1 Formatted Files	29 30
13	Script Files	10	27 Graphic User Interfaces	30
1 1	Products Division & D 177		28 Command Summary	31
14	Products, Division & Powers of Vectors 14.1 Scalar Product (*)	10 10 11 12 12		

1 MATLAB

- Matlab is an interactive system for doing numerical computations.
- A numerical analyst called Cleve Moler wrote the first version of Matlab in the 1970s. It has since evolved into a successful commercial software package.
- Matlab relieves you of a lot of the mundane tasks associated with solving problems numerically. This allows you to spend more time thinking, and encourages you to experiment.
- Matlab makes use of highly respected algorithms and hence you can be confident about your results.
- Powerful operations can be performed using just one or two commands.
- You can build up your own set of functions for a particular application.
- Excellent graphics facilities are available, and the pictures can be inserted into LATEX and Word documents.

These notes provide only a brief glimpse of the power and flexibility of the Matlab system. For a more comprehensive view we recommend the book

Matlab Guide D.J. Higham & N.J. Higham SIAM Philadelphia, 2000, ISBN: 0-89871-469-9.

2 Starting Up

2.1 Windows Systems

On Windows systems MATLAB is started by doubleclicking the MATLAB icon on the desktop or by selecting MATLAB from the start menu.

The starting procedure takes the user to the Command window where the Command line is indicated with '>>'. Used in the calculator mode all Matlab commands are entered to the command line from the keyboard.

Matlab can be used in a number of different ways or modes; as an advanced calculator in the calculator mode, in a high level programming language mode and as a subroutine called from a C-program. More information on the first two of these modes is given below.

Help and information on Matlab commands can be found in several ways,

• from the command line by using the 'help topic' command (see below),

- from the separate Help window found under the Help menu or
- from the Matlab helpdesk stored on disk or on a CD-ROM.

Another useful facility is to use the 'lookfor keyword' command, which searches the help files for the keyword. See Exercise 16.1 (page 17) for an example of its use.

2.2 Unix Systems

- You should have a directory reserved for saving files associated with Matlab. Create such a directory (mkdir) if you do not have one. Change into this directory (cd).
- Start up a new xterm window (do xterm & in the existing xterm window).
- Launch Matlab in one of the xterm windows with the command

matlab

After a short pause, the logo will be shown followed by

>>

where >> is the Matlab prompt.

Type quit at any time to exit from Matlab.

2.3 Command Line Help

Help is available from the command line prompt. Type help help for "help" (which gives a brief synopsis of the help system), help for a list of topics. The first few lines of this read

HELP topics:

```
matlab/general - General purpose commands.
matlab/ops - Operators and special char...
matlab/lang - Programming language const...
matlab/elmat - Elementary matrices and ma...
matlab/elfun - Elementary math functions.
matlab/specfun - Specialized math functions.
```

(truncated lines are shown with \dots). Then to obtain help on "Elementary math functions", for instance, type

>> help elfun

This gives rather a lot of information so, in order to see the information one screenful at a time, first issue the command more on, i.e.,

```
>> more on
>> help elfun
```

Hit any key to progress to the next page of information.

2.4 Demos

Demonstrations are invaluable since they give an indication of Matlabs capabilities. A comprehensive set are available by typing the command

>> demo

(Warning: this will clear the values of all current variables.)

3 Matlab as a Calculator

The basic arithmetic operators are + - * / ^ and these are used in conjunction with brackets: (). The symbol ^ is used to get exponents (powers): 2^4=16.

You should type in commands shown following the prompt: >>.

```
>> 2 + 3/4*5
ans =
5.7500
>>
```

Is this calculation 2 + 3/(4*5) or 2 + (3/4)*5? Matlab works according to the priorities:

- 1. quantities in brackets,
- 2. powers 2 + $3^2 \Rightarrow 2 + 9 = 11$,
- 3. * /, working left to right (3*4/5=12/5),
- 4. + -, working left to right (3+4-5=7-5),

Thus, the earlier calculation was for 2 + (3/4)*5 by priority 3.

4 Numbers & Formats

Matlab recognizes several different kinds of numbers

Type	Examples
Integer	1362, -217897
Real	1.234, -10.76
Complex	$3.21 - 4.3i (i = \sqrt{-1})$
Inf	Infinity (result of dividing by 0)
NaN	Not a Number, $0/0$

The "e" notation is used for very large or very small numbers:

```
-1.3412e+03 = -1.3412 \times 10^3 = -1341.2
-1.3412e-01 = -1.3412 \times 10^{-1} = -0.13412
```

All computations in MATLAB are done in double precision, which means about 15 significant figures. The format—how Matlab prints numbers—is controlled by the "format" command. Type help format for full list. Should you wish to switch back to the default format then format will suffice.

The command

format compact

is also useful in that it suppresses blank lines in the output thus allowing more information to be displayed.

Command	Example of Output
>>format short	31.4162(4-decimal places)
>>format short e	3.1416e+01
>>format long e	3.141592653589793e+01
>>format short	31.4162(4-decimal places)
>>format bank	31.42(2-decimal places)

5 Variables

```
>> 3-2^4
ans =
-13
>> ans*5
ans =
-65
```

The result of the first calculation is labelled "ans" by Matlab and is used in the second calculation where its value is changed.

We can use our own names to store numbers:

```
>> x = 3-2^4
x =
-13
>> y = x*5
y =
-65
```

so that \mathbf{x} has the value -13 and y=-65. These can be used in subsequent calculations. These are examples of **assignment statements**: values are assigned to variables. Each variable must be assigned a value before it may be used on the right of an assignment statement.

5.1 Variable Names

Legal names consist of any combination of letters and digits, starting with a letter. These are allowable:

These are **not** allowable:

Use names that reflect the values they represent.

Special names: you should avoid using

eps = 2.2204e-16 = 2^{-54} (The largest number such that 1 + eps is indistinguishable from 1) and pi = 3.14159... = π .

If you wish to do arithmetic with complex numbers, both i and j have the value $\sqrt{-1}$ unless you change them

```
>> i,j, i=3
ans = 0 + 1.0000i
ans = 0 + 1.0000i
i = 3
```

6 Suppressing output

One often does not want to see the result of intermediate calculations—terminate the assignment statement or expression with semi-colon

```
>> x=-13; y = 5*x, z = x^2+y
y =
-65
z =
104
>>
```

the value of x is hidden. Note also we can place several statements on one line, separated by commas or semicolons.

Exercise 6.1 In each case find the value of the expression in Matlab and explain precisely the order in which the calculation was performed.

7 Built-In Functions

7.1 Trigonometric Functions

Those known to Matlab are sin, cos, tan

and their arguments should be in radians.

e.g. to work out the coordinates of a point on a circle of radius 5 centred at the origin and having an elevation $30^o=\pi/6$ radians:

```
>> x = 5*cos(pi/6), y = 5*sin(pi/6)
x =
4.3301
y =
2.5000
```

The inverse trig functions are called asin, acos, atan (as opposed to the usual arcsin or sin^{-1} etc.). The result is in radians.

```
>> acos(x/5), asin(y/5)
ans = 0.5236
ans = 0.5236
>> pi/6
ans = 0.5236
```

7.2 Other Elementary Functions

These include sqrt, exp, log, log10

 $\exp(x)$ denotes the exponential function $\exp(x) = e^x$ and the inverse function is \log :

```
>> format long e, exp(log(9)), log(exp(9))
ans = 9.00000000000002e+00
ans = 9
>> format short
```

and we see a tiny rounding error in the first calculation. log10 gives logs to the base 10. A more complete list of elementary functions is given in Table 2 on page 32.

8 Vectors

These come in two flavours and we shall first describe **row vectors**: they are lists of numbers separated by either commas or spaces. The number of entries is known as the "length" of the vector and the entries are often referred to as "elements" or "components" of the vector. The entries must be enclosed in square brackets.

```
>> v = [ 1 3, sqrt(5)]
v =
1.0000 3.0000 2.2361
>> length(v)
ans =
3
```

Spaces can be vitally important:

We can do certain arithmetic operations with vectors of the same length, such as v and v3 in the previous section.

```
>> v + v3
ans =
              7.0000
    4.0000
                         7.2361
>> v4 = 3*v
v4 =
    3.0000
              9.0000
                         6.7082
>> v5 = 2*v - 3*v3
v5 =
   -7.0000
             -6.0000
                       -10.5279
>> v + v2
??? Error using ==> +
Matrix dimensions must agree.
```

i.e. the error is due to v and v2 having different lengths. A vector may be multiplied by a scalar (a number—see v4 above), or added/subtracted to another vector of the \mathbf{same} length. The operations are carried out elementwise.

We can build row vectors from existing ones:

```
>> w = [1 2 3], z = [8 9]
>> cd = [2*z,-w], sort(cd)
w =
1 2 3
z =
8 9
```

Notice the last command sort'ed the elements of cd into ascending order.

We can also change or look at the value of particular entries

8.1 The Colon Notation

This is a shortcut for producing row vectors:

More generally a:b:c produces a vector of entries starting with the value a, incrementing by the value b until it gets to c (it will not produce a value beyond c). This is why 1:-1 produced the empty vector [].

8.2 Extracting Bits of a Vector

To get the 3rd to 6th entries:

To get alternate entries:

What does r5(6:-2:1) give? See help colon for a fuller description.

8.3 Column Vectors

These have similar constructs to row vectors. When defining them, entries are separated by; or "newlines"

```
>> c = [ 1; 3; sqrt(5)]
    1.0000
    3.0000
    2.2361
>> c2 = [3
4
5]
c2
     3
     4
     5
>> c3 = 2*c - 3*c2
c3 =
   -7.0000
   -6.0000
  -10.5279
```

so column vectors may be added or subtracted **provided that they have the same length**.

8.4 Transposing

We can convert a row vector into a column vector (and vice versa) by a process called *transposing*—denoted by

```
>> w, w', c, c'
          -2
                  3
     1
ans
     1
    -2
     3
c =
    1.0000
    3.0000
    2.2361
ans =
               3.0000
                          2.2361
    1.0000
>> t = w + 2*c
t =
    3.0000
               4.0000
                          7.4721
>> T = 5*w'-2*c
T =
    3.0000
  -16.0000
   10.5279
```

If x is a complex vector, then x, gives the complex conjugate transpose of x:

```
>> x = [1+3i, 2-2i]
ans =
    1.0000 + 3.0000i    2.0000 - 2.0000i
>> x'
ans =
    1.0000 - 3.0000i
    2.0000 + 2.0000i
```

Note that the components of ${\tt x}$ were defined without a * operator; this means of defining complex numbers works even when the variable ${\tt i}$ already has a numeric value. To obtain the plain transpose of a complex number use .' as in

```
>> x.'
ans =
1.0000 + 3.0000i
2.0000 - 2.0000i
```

9 Keeping a record

Issuing the command

>> diary mysession

will cause all subsequent text that appears on the screen to be saved to the file mysession located in the directory in which Matlab was invoked. You may use any legal filename except the names on and off. The record may be terminated by

>> diary off

The file mysession may be edited with emacs to remove any mistakes.

If you wish to quit Matlab midway through a calculation so as to continue at a later stage:

>> save thissession

will save the current values of all variables to a file called thissession.mat. This file cannot be edited. When you next startup Matlab, type

>> load thissession

and the computation can be resumed where you left off. A list of variables used in the current session may be seen with

>> whos

See help whos and help save.

>> whos

Name	Size Elemen	ts Bytes	Density	Complex
ans	1 by 1 1	8	Full	No
v	1 by 3 3	24	Full	No
v1	1 by 2 2	16	Full	No
v2	1 by 2 2	16	Full	No
v3	1 by 3 3	24	Full	No
v4	1 by 3 3	24	Full	No
x	1 by 1 1	8	Full	No
У	1 by 1 1	8	Full	No

Grand total is 16 elements using 128 bytes

10 Plotting Elementary Functions

Suppose we wish to plot a graph of $y = \sin 3\pi x$ for $0 \le x \le 1$. We do this by sampling the function at a sufficiently large number of points and then joining up the points (x, y) by straight lines. Suppose we take N+1 points equally spaced a distance h apart:

$$>> N = 10; h = 1/N; x = 0:h:1;$$

defines the set of points x = 0, h, 2h, ..., 1 - h, 1. The corresponding y values are computed by

$$>> y = sin(3*pi*x);$$

and finally, we can plot the points with

The result is shown in Figure 1, where it is clear that the value of N is too small.

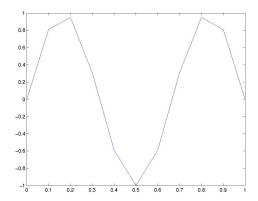


Figure 1: Graph of $y = \sin 3\pi x$ for $0 \le x \le 1$ using h = 0.1.

On changing the value of N to 100:

```
>> N = 100; h = 1/N; x = 0:h:1;
>> y = \sin(3*pi*x); plot(x,y)
```

we get the picture shown in Figure 2.

10.1 Plotting—Titles & Labels

To put a title and label the axes, we use

The *strings* enclosed in single quotes, can be anything of our choosing. Some simple LATEX commands are available for formatting mathematical expressions and Greek characters—see Section 10.9.

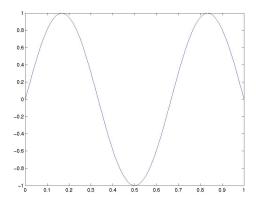


Figure 2: Graph of $y = \sin 3\pi x$ for $0 \le x \le 1$ using h = 0.01.

10.2 Grids

A dotted grid may be added by

>> grid

This can be removed using either grid again, or grid off.

10.3 Line Styles & Colours

The default is to plot solid lines. A solid white line is produced by

The third argument is a string whose first character specifies the colour(optional) and the second the line style. The options for colours and styles are:

	$\operatorname{Colours}$	Li	ne Styles
У	yellow		point
m	magenta	0	circle
С	cyan	x	x-mark
r	red	+	$_{ m plus}$
g	green	-	solid
b	blue	*	star
W	white	:	dotted
k	black		dashdot
			dashed

The number of available plot symbols is wider than shown in this table. Use help plot to obtain a full list. See also help shapes.

10.4 Multi-plots

Several graphs may be drawn on the same figure as in

A descriptive legend may be included with

>> legend('Sin curve', 'Cos curve')

which will give a list of line—styles, as they appeared in the plot command, followed by a brief description. Matlab fits the legend in a suitable position, so as not to conceal the graphs whenever possible.

For further information do help plot etc.

The result of the commands

```
>> plot(x,y,'w-',x,cos(3*pi*x),'g--')
>> legend('Sin curve','Cos curve')
>> title('Multi-plot ')
>> xlabel('x axis'), ylabel('y axis')
>> grid
```

is shown in Figure 3. The legend may be moved manually by dragging it with the mouse.

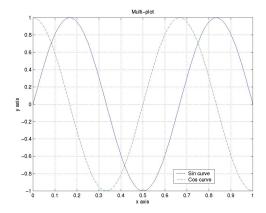


Figure 3: Graph of $y = \sin 3\pi x$ and $y = \cos 3\pi x$ for $0 \le x \le 1$ using h = 0.01.

10.5 Hold

A call to plot clears the graphics window before plotting the current graph. This is not convenient if we wish to add further graphics to the figure at some later stage. To stop the window being cleared:

"hold on" holds the current picture; "hold off" releases it (but does not clear the window, which can be done with clf). "hold" on its own toggles the hold state.

10.6 Hard Copy

To obtain a printed copy select Print from the File menu on the Figure toolbar.

Alternatively one can save a figure to a file for later printing (or editing). A number of formats is available (use help print to obtain a list). To save a file in "Encapsulated PostScript" format, issue the Matlab command

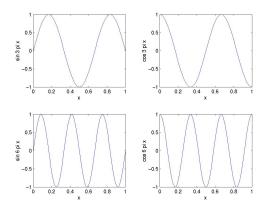
which will save a copy of the image in a file called fig1.eps.

10.7 Subplot

The graphics window may be split into an $m \times n$ array of smaller windows into which we may plot one or more graphs. The windows are counted 1 to mn row—wise, starting from the top left. Both hold and grid work on the current subplot.

```
>> subplot(221), plot(x,y)
>> xlabel('x'),ylabel('sin 3 pi x')
>> subplot(222), plot(x,cos(3*pi*x))
>> xlabel('x'),ylabel('cos 3 pi x')
>> subplot(223), plot(x,sin(6*pi*x))
>> xlabel('x'),ylabel('sin 6 pi x')
>> subplot(224), plot(x,cos(6*pi*x))
>> xlabel('x'),ylabel('cos 6 pi x')
```

subplot(221) (or subplot(2,2,1)) specifies that the window should be split into a 2×2 array and we select the first subwindow.



10.8 Zooming

We often need to "zoom in" on some portion of a plot in order to see more detail. This is easily achieved using the command

>> zoom

Pointing the mouse to the relevant position on the plot and clicking the left mouse button will zoom in by a factor of two. This may be repeated to any desired level.

Clicking the right mouse button will zoom out by a factor of two

Holding down the left mouse button and dragging the mouse will cause a rectangle to be outlined. Releasing the button causes the contents of the rectangle to fill the window.

zoom off turns off the zoom capability.

Exercise 10.1 Draw graphs of the functions

$$y = \cos x$$

 $y = x$

for $0 \le x \le 2$ on the same window. Use the zoom facility to determine the point of intersection of the two

curves (and, hence, the root of $x = \cos x$) to two significant figures.

The command clf clears the current figure while close 1 will close the window labelled "Figure 1". To open a new figure window type figure or, to get a window labelled "Figure 9", for instance, type figure (9). If "Figure 9" already exists, this command will bring this window to the foreground and the result subsequent plotting commands will be drawn on it.

10.9 Formatted text on Plots

It is possible to change to format of text on plots so as to increase or decrease its size and also to typeset simple mathematical expressions (in LATEX form).

We shall give two illustrations.

First we plot the first 100 terms in the sequence $\{x_n\}$ given by $x_n = \left(1 + \frac{1}{n}\right)^n$ and then graph the function $\phi(x) = x^3 \sin^2(3\pi x)$ on the interval $-1 \le x \le 1$. The commands

produce the graph shown below. The salient features of these commands are

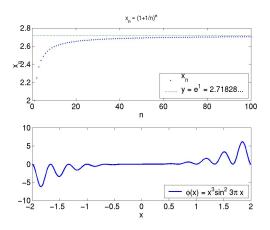
- The first line increases the size of the default font size used for the axis labels, legends and titles.
- 2. The size of the plot symbol "." is changed from the default (6) to size 8 by the additional string followed by value "'markersize',8".
- The strings 'x_n' are formatted as x_n to give subscripts while 'x^3 leads to superscripts x³.
 Note also that sin²3πx translates into the Matlab command sin(3*pi*x).^2—the position of the exponent is different.
- 4. Greek characters $\alpha, \beta, \ldots, \omega, \Omega$ are produced by the strings '\alpha', '\beta', \ldots, '\omega', '\omega'. the integral symbol: \int is produced by '\int'.
- 5. The thickness of the line used in the lower graph is changed from its default value (0.5) to 2.
- 6. Use help legend to determine the meaning of the last argument in the legend commands.

One can determine the current value of any plot property by first obtaining its "handle number" and then using the get command such as

```
>> handle = plot (x,y,'.')
>> get (handle,'markersize')
ans =
6
```

Experiment also with set (handle) (which will list possible values for each property) and set (handle, 'mark Also, all plot properties can be edited from the Figure window by selecting the Tools menu from the toolbar. For instance, to change the linewidth of a graph, first select the curve by double clicking (it should then change its appearance) and then select Line Properties...

from the <u>Tools</u>. This will pop up a dialogue window from which the width, colour, style,... of the curve may be changed.



10.10 Controlling Axes

Once a plot has been created in the graphics window you may wish to change the range of x and y values shown on the picture.

```
>> clf, N = 100; h = 1/N; x = 0:h:1;
>> y = sin(3*pi*x); plot(x,y)
>> axis([-0.5 1.5 -1.2 1.2]), grid
```

The axis command has four parameters, the first two are the minimum and maximum values of x to use on the axis and the last two are the minimum and maximum values of y. Note the square brackets. The result of these commands is shown in Figure 4. Look at help axis and experiment with the commands axis equal, axis verb, axis square, axis normal, axis tight in any order.

11 Keyboard Accelerators

One can recall previous Matlab commands by using the \uparrow and \downarrow cursor keys. Repeatedly pressing \uparrow will review the previous commands (most recent first) and, if you want to re-execute the command, simply press the return key.

To recall the most recent command starting with p, say, type p at the prompt followed by \uparrow . Similarly, typing

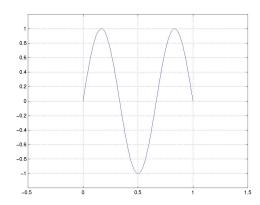


Figure 4: The effect of changing the axes of a plot.

pr followed by \uparrow will recall the most recent command starting with pr.

Once a command has been recalled, it may be edited (changed). You can use \leftarrow and \rightarrow to move backwards and forwards through the line, characters may be inserted by typing at the current cursor position or deleted using the Del key. This is most commonly used when long command lines have been mistyped or when you want to re—execute a command that is very similar to one used previously.

The following emacs-like commands may also be used:

```
cntrl a move to start of line
cntrl e move to end of line
cntrl f move forwards one character
cntrl b move backwards one character
cntrl d delete character under the cursor
```

Once you have the command in the required form, press

Exercise 11.1 Type in the commands

```
>> x = -1:0.1:1;
>> plot(x,sin(pi*x),'w-')
>> hold on
>> plot(x,cos(pi*x),'r-')
```

Now use the cursor keys with suitable editing to execute:

```
>> x = -1:0.05:1;
>> plot(x,sin(2*pi*x),'w-')
>> plot(x,cos(2*pi*x),'r-.'), hold off
```

12 Copying to and from Word and other applications

There are many situations where one wants to copy the output resulting from a Matlab command (or commands) into a Windows application such as Word or into a Unix file editor such as "emacs" or "vi".

12.1 Window Systems

Copying material is made possible on the Windows operating system by using the Windows clipboard.

Also, pictures can be exported to files in a number of alternative formats such as encapsulated postscript format or in jpeg format. Matlab is so frequently used as an analysis tool that many manufacturers of measurement systems and software find it convenient to provide interfaces to Matlab which make it possible, for instance, to import measured data directly into a *.mat Matlab file (see load and save in Section 9).

Example 12.1 Copying a figure into Word.

Diagrams prepared in Matlab are easily exported to other Windows applications such as Word. Suppose a plot of the functions $\sin(2\pi ft)$ and $\sin(2\pi ft + \pi/4)$, with f=100, is needed in a report written in Word. We create a time vector, t, with 500 points distributed over 5 periods and then evaluate and plot the two function vectors.

```
>> t = [1:1:500]/500/20;
>> f = 100;
>> y1 = sin(2*pi*f*t);
>> y2 = sin(2*pi*f*t+pi/4);
>> plot(t,y1,'-',t,y2,'--');
>> axis([0 0.05 -1.5 1.5]);
>> grid
```

In order to copy the plot into a Word document

- Select "Copy Figure" under the Edit menu on the figure windows toolbar.
- Switch to the Word application if it is already running, otherwise open a Word document.
- Place the cursor in the desired position in the document and select "Paste" under the "Edit" menu in the Word tool bar.

12.2 Unix Systems

In order to carry out the following exercise, you should have Matlab running in one window and either Emacs or Vi running in another.

To copy material from one window to another, (here L means click Left Mouse Button, etc)

First select the material to copy by L on the start of the material you want and then either dragging the mouse (with the buttom down) to highlight the text, or R at the end of the material. Next move the mouse into the other window and L at the location you want the text to appear. Finally, click the M.

When copying from another application into Matlab you can only copy material to the prompt line. On Unix systems figures are normally saved in files (see Section 10.6) which are then imported into other documents.

13 Script Files

Script files are normal ASCII (text) files that contain Matlab commands. It is essential that such files have names having an extension .m (e.g., Exercise4.m) and, for this reason, they are commonly known as *m-files*. The commands in this file may then be executed using >> Exercise4

Note: the command does not include the file name extension .m.

It is only the output from the commands (and not the commands themselves) that are displayed on the screen. Script files are created with your favourite editor under Unix while, under Windows, click on the "New Document" icon at the top left of the main Matlab window to pop up a new window showing the "M-file Editor". Type in your commands and then save (to a file with a .m extension).

To see the commands in the command window prior to their execution:

>> echo on

and echo off will turn echoing off.

Any text that follows % on a line is ignored. The main purpose of this facility is to enable comments to be included in the file to describe its purpose.

To see what m-files you have in your current directory, use

>> what

Exercise 13.1 1. Type in the commands from §10.7 into a file called exsub.m.

- 2. Use what to check that the file is in the correct area.
- 3. Use the command type exsub to see the contents of the file.
- 4. Execute these commands.

See §21 for the related topic of function files.

14 Products, Division & Powers of Vectors

14.1 Scalar Product (*)

We shall describe two ways in which a meaning may be attributed to the product of two vectors. In both cases the vectors concerned must have the same length.

The first product is the standard scalar product. Suppose that \underline{u} and \underline{v} are two vectors of length n, \underline{u} being a **row** vector and \underline{v} a **column** vector:

$$\underline{u} = [u_1, u_2, \dots, u_n], \qquad \underline{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}.$$

The scalar product is defined by multiplying the corresponding elements together and adding the results to give a single number (scalar).

$$\underline{u}\,\underline{v} = \sum_{i=1}^{n} u_i v_i.$$

For example, if
$$\underline{u}=[10,-11,12],$$
 and $\underline{v}=\begin{bmatrix} 20\\ -21\\ -22 \end{bmatrix}$

then n=3 and

$$\underline{u}\,\underline{v} = 10 \times 20 + (-11) \times (-21) + 12 \times (-22) = 167.$$

We can perform this product in Matlab by

>> u = [10, -11, 12], v = [20; -21; -22]
>> prod =
$$u*v$$
 % row times column vector

Suppose we also define a row vector ${\bf w}$ and a column vector ${\bf z}$ by

and we wish to form the scalar products of \boldsymbol{u} with \boldsymbol{w} and \boldsymbol{v} with $\boldsymbol{z}.$

>> u*w
??? Error using ==> *
Inner matrix dimensions must agree.

an error results because w is not a column vector. Recall from page 5 that transposing (with ') turns column vectors into row vectors and vice versa.

So, to form the scalar product of two row vectors or two column vectors,

We shall refer to the Euclidean length of a vector as the **norm** of a vector; it is denoted by the symbol $\|\underline{u}\|$ and defined by

$$\|\underline{u}\| = \sqrt{\sum_{i=1}^{n} |u_i|^2},$$

where n is its dimension. This can be computed in Matlab in one of two ways:

where norm is a built—in Matlab function that accepts a vector as input and delivers a scalar as output. It can also be used to compute other norms: help norm.

Exercise 14.1 The angle, θ , between two column vectors \underline{x} and y is defined by

$$\cos \theta = \frac{\underline{x}' \underline{y}}{\|\underline{x}\| \|\underline{y}\|}.$$

Use this formula to determine the cosine of the angle between

$$\underline{x} = [1, 2, 3]'$$
 and $y = [3, 2, 1]'$.

Hence find the angle in degrees.

14.2 Dot Product (.*)

The second way of forming the product of two vectors of the same length is known as the Hadamard product. It is not often used in Mathematics but is an invaluable Matlab feature. It involves vectors of the same type. If \underline{u} and \underline{v} are two vectors of the same type (both row vectors or both column vectors), the mathematical definition of this product, which we shall call the **dot product**, is the **vector** having the components

$$\underline{u} \cdot \underline{v} = [u_1 v_1, u_2 v_2, \dots, u_n v_n].$$

The result is a vector of the same length and type as \underline{u} and \underline{v} . Thus, we simply multiply the corresponding elements of two vectors.

In Matlab, the product is computed with the operator .* and, using the vectors u, v, w, z defined on page 11,

Example 14.1 Tabulate the function $y = x \sin \pi x$ for $x = 0, 0.25, \dots, 1$.

It is easier to deal with column vectors so we first define a vector of x-values: (see Transposing: $\S 8.4$)

$$>> x = (0:0.25:1)$$
;

To evaluate y we have to multiply each element of the vector x by the corresponding element of the vector $\sin \pi x$:

x	X	$\sin \pi x = 3$	$x\sin\pi x$
0	×	0 =	0
0.2500	×	0.7071 =	0.1768
0.5000	×	1.0000 =	0.5000
0.7500	×	0.7071 =	0.5303
1.0000	×	0.0000 =	0.0000

To carry this out in Matlab:

Note: a) the use of pi, b) x and sin(pi*x) are both column vectors (the sin function is applied to each element of the vector). Thus, the dot product of these is also a column vector.

14.3 Dot Division of Arrays (./)

There is no mathematical definition for the division of one vector by another. However, in Matlab, the operator ./ is defined to give element by element division—it is therefore only defined for vectors of the same size and type.

The previous calculation required division by 0—notice the Inf, denoting infinity, in the answer.

Warning: Divide by zero ans = 9 10 NaN 12 13

Here we are warned about 0/0—giving a NaN (Not a Number).

Example 14.2 Estimate the limit

$$\lim_{x \to 0} \frac{\sin \pi x}{x}.$$

The idea is to observe the behaviour of the ratio $\frac{\sin \pi x}{x}$ for a sequence of values of x that approach zero. Suppose that we choose the sequence defined by the column vector

```
>> x = [0.1; 0.01; 0.001; 0.0001]then
```

```
>> sin(pi*x)./x
ans =
3.0902
3.1411
3.1416
3.1416
```

which suggests that the values approach π . To get a better impression, we subtract the value of π from each entry in the output and, to display more decimal places, we change the format

```
>> format long
>> ans -pi
ans =
-0.05142270984032
-0.00051674577696
-0.00000516771023
-0.00000005167713
```

Can you explain the pattern revealed in these numbers? We also need to use ./ to compute a scalar divided by a vector:

```
>> 1/x
??? Error using ==> /
Matrix dimensions must agree.
>> 1./x
ans =
10 100 1000 10000
so 1./x works, but 1/x does not.
```

14.4 Dot Power of Arrays (.^)

To square each of the elements of a vector we could, for example, do u.*u. However, a neater way is to use the .^ operator:

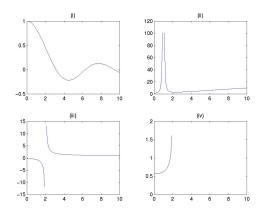
```
>> u.^2
ans =
   100
          121
                144
>> u.*u
ans =
   100
          121
                144
>> u.^4
ans
                     14641
                                  20736
       10000
>> v.^2
   400
   441
   484
>> u.*w.^(-2)
ans =
    2.5000 -11.0000
                          1.3333
```

Recall that powers (. ^ in this case) are done first, before any other arithmetic operation.

15 Examples in Plotting

Example 15.1 Draw graphs of the functions

i)
$$y = \frac{\sin x}{x}$$
 ii) $u = \frac{1}{(x-1)^2} + x$
iii) $v = \frac{x^2+1}{x^2-4}$ iv) $w = \frac{(10-x)^{1/3}-(10-x)^{1/3}}{(4-x^2)^{1/2}}$



Note the repeated use of the "dot" operators. Experiment by changing the axes (page 9), grids (page 7) and hold(page 7).

- >> grid
- >> grid
- >> hold on
- >> plot(x,v,'--'), hold off, plot(x,y,':')

Exercise 15.1 Enter the vectors

$$\underline{U} = \begin{bmatrix} 6, 2, 4 \end{bmatrix}, \quad \underline{V} = \begin{bmatrix} 3, -2, 3, 0 \end{bmatrix},$$

$$\underline{W} = \begin{bmatrix} 3 \\ -4 \\ 2 \\ -6 \end{bmatrix}, \quad \underline{Z} = \begin{bmatrix} 3 \\ 2 \\ 2 \\ 7 \end{bmatrix}$$

into Matlab.

1. Which of the products

is legal? State whether the legal products are row or column vectors and give the values of the legal results. 2. Tabulate the functions

$$y = (x^2 + 3)\sin \pi x^2$$

and

$$z = \sin^2 \pi x / (x^{-2} + 3)$$

for $x = 0, 0.2, \dots, 10$. Hence, tabulate the function

$$w = \frac{(x^2 + 3)\sin \pi x^2 \sin^2 \pi x}{(x^{-2} + 3)}$$

Plot a graph of w over the range $0 \le x \le 10$.

16 Matrices—Two–Dimensional Arrays

Row and Column vectors are special cases of **matrices**. An $m \times n$ matrix is a rectangular array of numbers having m rows and n columns. It is usual in a mathematical setting to include the matrix in either round or square brackets—we shall use square ones. For example, when m=2, n=3 we have a 2×3 matrix such

$$A = \left[\begin{array}{ccc} 5 & 7 & 9 \\ 1 & -3 & -7 \end{array} \right]$$

To enter such an matrix into Matlab we type it in row by row using the same syntax as for vectors:

 $>> B = [-1 \ 2 \ 5; \ 9 \ 0 \ 5]$

Rows may be separated by semi-colons rather than a new line:

So A and B are 2×3 matrices, C is 3×2 and D is 3×5 . In this context, a row vector is a $1\times n$ matrix and a column vector a $m\times 1$ matrix.

16.1 Size of a matrix

We can get the size (dimensions) of a matrix with the command ${\tt size}$

So A is 2×3 and x is 3×1 (a column vector). The last command size(ans) shows that the *value* returned by size is itself a 1×2 matrix (a row vector). We can save the results for use in subsequent calculations.

16.2 Transpose of a matrix

Transposing a vector changes it from a row to a column vector and vice versa (see §8.4). The extension of this idea to matrices is that transposing interchanges rows with the corresponding columns: the 1st row becomes the 1st column, and so on.

16.3 Special Matrices

Matlab provides a number of useful built—in matrices of any desired size.

ones(m,n) gives an $m \times n$ matrix of 1's,

zeros(m,n) gives an $m \times n$ matrix of 0's,

The second command illustrates how we can construct a matrix based on the size of an existing one. Try ones(size(D)).

An $n \times n$ matrix that has the same number of rows and columns and is called a **square** matrix.

A matrix is said to be **symmetric** if it is equal to its transpose (i.e. it is unchanged by transposition):

16.4 The Identity Matrix

The $n \times n$ identity matrix is a matrix of zeros except for having ones along its leading diagonal (top left to bottom right). This is called eye(n) in Matlab (since mathematically it is usually denoted by I).

Notice that multiplying the 3×1 vector **x** by the 3×3 identity **I** has no effect (it is like multiplying a number by 1).

16.5 Diagonal Matrices

A diagonal matrix is similar to the identity matrix except that its diagonal entries are not necessarily equal to 1.

$$D = \left[\begin{array}{rrr} -3 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 2 \end{array} \right]$$

is a 3×3 diagonal matrix. To construct this in Matlab, we could either type it in directly

but this becomes impractical when the dimension is large (e.g. a 100×100 diagonal matrix). We then use the diag function.We first define a vector d, say, containing the values of the diagonal entries (in order) then diag(d) gives the required matrix.

On the other hand, if A is any matrix, the command diag(A) extracts its diagonal entries:

Notice that the matrix does not have to be square.

16.6 Building Matrices

It is often convenient to build large matrices from smaller ones:

so we have added an extra column (x) to C in order to form G and have stacked A and B on top of each other to form H.

```
>> J = [1:4; 5:8; 9:12; 20 0 5 4]
            2
                    3
                           4
      1
     5
            6
                    7
                           8
     9
            10
                   11
                          12
    20
            0
                    5
                           4
>> K = [ diag(1:4) J; J' zeros(4,4)]
                           2
                                3
   1
        0
            0
                  0
                      1
   0
        2
             0
                 0
                      5
                           6
                                7
                                     8
   0
        0
             3
                 0
                      9
                          10
                               11
                                   12
   0
        0
            0
                 4
                     20
                           0
                                5
        5
            9
                                0
   1
                20
                      0
                           0
                                     0
   2
        6
           10
                      0
                           0
                                0
                                     0
                 0
        7
   3
                 5
                      0
                           0
                                0
                                     0
           11
        8
                      0
                           0
                                0
                                     0
```

The command spy(K) will produce a graphical display of the location of the nonzero entries in K (it will also give a value for nz—the number of nonzero entries):

>> x = 0:0.1:0.5;

16.7 Tabulating Functions

This has been addressed in earlier sections but we are now in a position to produce a more suitable table format.

Example 16.1 Tabulate the functions $y = 4 \sin 3x$ and $u = 3 \sin 4x$ for $x = 0, 0.1, 0.2, \dots, 0.5$.

```
>> y = 4*sin(3*x); u = 3*sin(4*x);
>> [ x' y' u']
ans =
         0
                    0
                               0
    0.1000
              1.1821
                         1.1683
    0.2000
              2.2586
                         2.1521
    0.3000
              3.1333
                         2.7961
    0.4000
               3.7282
                         2.9987
    0.5000
              3.9900
                         2.7279
```

Note the use of transpose (') to get column vectors. (we could replace the last command by [x; y; u;]') We could also have done this more directly:

```
>> x = (0:0.1:0.5)';
>> [x 4*sin(3*x) 3*sin(4*x)]
```

16.8 Extracting Bits of Matrices

We may extract sections from a matrix in much the same way as for a vector (page 5).

Each element of a matrix is indexed according to which row and column it belongs to. The entry in the *i*th row and *j*th column is denoted mathematically by $A_{i,j}$ and, in Matlab, by A(i,j). So

```
>> J
J =
1 2 3 4
5 6 7 8
```

In the following examples we extract i) the 3rd column, ii) the 2nd and 3rd columns, iii) the 4th row, and iv) the "central" 2×2 matrix. See §8.1.

Thus, : on its own refers to the entire column or row depending on whether it is the first or the second index.

16.9 Dot product of matrices (.*)

The dot product works as for vectors: corresponding elements are multiplied together—so the matrices involved must have the same size.

16.10 Matrix-vector products

We turn next to the definition of the product of a matrix with a vector. This product is only defined for **column vectors** that have the same number of entries as the matrix has columns. So, if A is an $m \times n$ matrix and \underline{x} is a column vector of length n, then the matrix–vector Ax is legal.

An $m \times n$ matrix times an $n \times 1$ matrix \Rightarrow a $m \times 1$ matrix.

We visualise A as being made up of m row vectors stacked on top of each other, then the product corresponds to taking the **scalar** product (See §14.1) of each row of A with the vector \underline{x} : The result is a column vector with m entries.

$$A\underline{x} = \begin{bmatrix} 5 & 7 & 9 \\ 1 & -3 & -7 \end{bmatrix} \begin{bmatrix} 8 \\ -4 \\ 1 \end{bmatrix}$$
$$= \begin{bmatrix} 5 \times 8 + 7 \times (-4) + 9 \times 1 \\ 1 \times 8 + (-3) \times (-4) + (-7) \times 1 \end{bmatrix}$$
$$= \begin{bmatrix} 21 \\ 13 \end{bmatrix}$$

It is somewhat easier in Matlab:

>> A = [5 7 9; 1 -3 -7]
A =

5 7 9
1 -3 -7

>> x = [8; -4; 1]
x =

8 -4
1
>> A*x
ans =
21
13

$$(m \times n) \text{ times } (n \times 1) \Rightarrow (m \times 1).$$

??? Error using ==> *
Inner matrix dimensions must agree.

Unlike multiplication in arithmetic, A*x is not the same as x*A.

16.11 Matrix-Matrix Products

To form the product of an $m \times n$ matrix A and a $n \times p$ matrix B, written as AB, we visualise the first matrix (A) as being composed of m row vectors of length n stacked on top of each other while the second (B) is visualised as being made up of p column vectors of length n:

$$A = m \text{ rows} \left\{ \begin{bmatrix} \vdots \\ \vdots \\ \end{bmatrix}, B = \underbrace{\begin{bmatrix} \vdots \\ \end{bmatrix}}_{p \text{ columns}} \right\}$$

The entry in the *i*th row and *j*th column of the product is then the **scalar** product of the *i*th row of A with the *j*th column of B. The product is an $m \times p$ matrix:

$$(m \times \boxed{n) \text{ times } (n \times p)} \Rightarrow (m \times p).$$

Check that you understand what is meant by working out the following examples by hand and comparing with the Matlab answers.

We see that E = C' suggesting that

$$(A*B)' = B*A'$$

Why is $B*A$ a 3×3 matrix while $A*B$ is 2×2 ?

Exercise 16.1 It is often necessary to factorize a matrix, e.g., A = BC or $A = S^TXS$ where the factors are required to have specific properties. Use the 'lookfor keyword' command to make a list of factorizations commands in Matlab.

16.12 Sparse Matrices

Matlab has powerful techniques for handling sparse matrices — these are generally large matrices (to make the

extra work involved worthwhile) that have only a very small proportion of non–zero entries.

Example 16.2 Create a sparse 5×4 matrix S having only 3 non-zero values: $S_{1,2} = 10$, $S_{3,3} = 11$ and $S_{5,4} = 12$.

We first create 3 vectors containing the i-index, the j-index and the corresponding values of each term and we then use the sparse command.

The matrix T is a "full" version of the sparse matrix S.

Example 16.3 Develop Matlab code to create, for any given value of n, the sparse (tridiagonal) matrix

$$B = \begin{bmatrix} 1 & n & & & & & & \\ -2 & 2 & n-1 & & & & & \\ & -3 & 3 & n-2 & & & & \\ & & \ddots & \ddots & \ddots & \ddots & \\ & & & -n+1 & n-1 & 1 \\ & & & & -n & n \end{bmatrix}$$

We define three COLUMN vectors, one for each "diagonal" of non–zeros and then assemble the matrix using spdiags (short for sparse diagonals). The vectors are named 1, d and u. They must all have the same length and only the first n-1 terms of 1 are used while the last n-1 terms of u are used. spdiags places these vectors in the diagonals labelled -1, 0 and 1 (0 defers to the leading diagonal, negatively numbered diagonals lie below the leading diagonal, etc.)

```
>> n = 5;
>> l = -(2:n+1)'; d = (1:n )'; u = ((n+1):-1:2)';
>> B = spdiags([l' d' u'],-1:1,n,n);
>> full(B)
ans =

    1     5     0     0     0
    -2     2     4     0     0
    0     -3     3     3     0
    0     0     -4     4     2
    0     0     0     -5     5
```

17 Systems of Linear Equations

Mathematical formulations of engineering problems often lead to sets of simultaneous linear equations. Such is the case, for instance, when using the finite element method (FEM).

A general system of linear equations can be expressed in terms of a coefficient matrix A, a right-hand-side (column) vector \mathbf{b} and an unknown (column) vector \mathbf{x}

$$A\mathbf{x} = \mathbf{b}$$

or, componentwise, as

$$\begin{array}{lll} a_{1,1}x_1 + a_{1,2}x_2 + \cdots a_{1,n}x_n & = b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots a_{2,n}x_n & = b_2 \\ & \vdots \\ a_{n,1}x_1 + a_{n,2}x_2 + \cdots a_{n,n}x_n & = b_n \end{array}$$

When A is non-singular and square $(n \times n)$, meaning that the number of *independent* equations is equal to the number of unknowns, the system has a unique solution given by

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

where A^{-1} is the inverse of A. Thus, the solution vector \mathbf{x} can, in principle, be calculated by taking the inverse of the coefficient matrix A and multiplying it on the right with the right-hand-side vector \mathbf{b} .

This approach based on the matrix inverse, though formally correct, is at best inefficient for practical applications (where the number of equations may be extremely large) but may also give rise to large numerical errors unless appropriate techniques are used. These issues are discussed in most courses and texts on numerical methods. Various stable and efficient solution techniques have been developed for solving linear equations and the most appropriate in any situation will depend on the properties of the coefficient matrix A. For instance, on whether or not it is symmetric, or positive definite or if it has a particular structure (sparse or full). Matlab is equipped with many of these special techniques in its routine library and they are invoked automatically.

The standard Matlab routine for solving systems of linear equations is invoked by calling the matrix leftdivision routine,

$$>> x = A \setminus b$$

where "\" is the matrix left-division operator known as "backslash" (see help backslash).

Exercise 17.1 Enter the symmetric coefficient matrix and right-hand-side vector **b** given by

$$A = \begin{bmatrix} 2 & -1 & 0 \\ 1 & -2 & 1 \\ 0 & -1 & 2 \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

and solve the system of equations $A\mathbf{x} = \mathbf{b}$ using the three alternative methods:

- i) $\mathbf{x} = A^{-1}\mathbf{b}$, (the inverse A^{-1} may be computed in Matlab using inv(A).)
- $ii) x = A \setminus b$,
- iii) $\mathbf{x}^T = \mathbf{b}^t A^T$ leading to $\mathbf{x}T = \mathbf{b}$, / A which makes use of the "slash" or "right division" operator "/". The required solution is then the transpose of the row vector $\mathbf{x}T$.

Exercise 17.2 Use the backslash operator to solve the complex system of equations for which

$$A = \begin{bmatrix} 2+2i & -1 & 0 \\ -1 & 2-2i & -1 \\ 0 & -1 & 2 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1+i \\ 0 \\ 1-i \end{bmatrix}$$

Exercise 17.3 Find information on the matrix inversion command 'inv' using each of the methods listed in Section 2 for obtaining help.

What kind of matrices are the 'inv' command applicable to?

Obviously problems may occur if the inverted matrix is nearly singular. Suggest a command that can be used to give an indication on whether the matrix is nearly singular or not. [Hint: see the topics referred to by 'help inv'.]

17.1 Overdetermined system of linear equations

An overdetermined system of linear equations is a one with more equations (m) than unknowns (n), i.e., the coefficient matrix has more rows than columns (m > n). Overdetermined systems frequently appear in mathematical modelling when the parameters of a model are determined by fitting to experimental data. Formally the system looks the same as for square systems but the coefficient matrix is rectangular and so it is not possible to compute an inverse. In these cases a solution can be found by requiring that the magnitude of the residual vector \mathbf{r} , defined by

$$\mathbf{r} = A\mathbf{x} - \mathbf{b},$$

be minimized. The simplest and most frequently used measure of the magnitude of ${\bf r}$ is require the Euclidean length (or norm—see Section 14.1) which corresponds to the sum of squares of the components of the residual. This approach leads to the least squares solution of the overdetermined system. Hence the least squares solution is defined as the vector ${\bf x}$ that minimizes

$$\mathbf{r}^T \mathbf{r}$$

It may be shown that the required solution satisfies the so-called "normal equations"

$$C\mathbf{x} = \mathbf{d}$$
, where $C = A^T A$ and $\mathbf{d} = A^T \mathbf{b}$.

This system is well–known that the solution of this system can be overwhelmed by numerical rounding error in practice unless great care is taken in its solution (a large part of the difficulty is inherent in computing the matrix–matrix product A^TA). As in the solution of square systems of linear equations, special techniques have been developed to address these issues and they have been incorporated into the Matlab routine library. This means that a direct solution to the problem of overdetermined equations is available in Matlab through its left division operator "\". When the matrix A is not square, the operation

$$x = A b$$

automatically gives the least squares solution to $A\mathbf{x} = \mathbf{b}$. This is illustrated in the next example.

Example 17.1 A spring is a mechanical element which, and the least squares solution to this system is given by for the simplest model, is characterized by a linear forcedeformation relationship

$$F = kx$$

F being the force loading the spring, k the spring constant or stiffness and x the spring deformation. In reality the linear force-deformation relationship is only an approximation, valid for small forces and deformations. A more accurate relationship, valid for larger deformations, is obtained if non-linear terms are taken into account. Suppose a spring model with a quadratic relationship

$$F = k_1 x + k_2 x^2$$

is to be used and that the model parameters, k_1 and k_2 , are to be determined from experimental data. Five independent measurements of the force and the corresponding spring deformations are measured and these are presented in Table 1.

Force F [N]	Deformation x [cm]
5	0.001
50	0.011
500	0.013
1000	0.30
2000	0.75

Table 1: Measured force-deformation data for spring.

Using the quadratic force-deformation relationship together with the experimental data yields an overdetermined system of linear equations and the components of the residual are given by

$$\begin{array}{lll} r_1 & = x_1k_1 + x_1^2k_2 - F_1 \\ r_2 & = x_2k_1 + x_2^2k_2 - F_2 \\ r_3 & = x_3k_1 + x_3^2k_2 - F_3 \\ r_4 & = x_4k_1 + x_4^2k_2 - F_4 \\ r_5 & = x_5k_1 + x_5^2k_2 - F_5. \end{array}$$

These lead to the matrix and vector definitions

$$A = \begin{bmatrix} x_1 & x_1^2 \\ x_2 & x_2^2 \\ x_3 & x_3^2 \\ x_4 & x_4^2 \\ x_5 & x_5^2 \end{bmatrix} \text{ and } \mathbf{b} = \begin{bmatrix} F_1 \\ F_2 \\ F_3 \\ F_4 \\ F_5. \end{bmatrix}$$

The appropriate Matlab commands give (the components of x are all multiplied by 1e-2, i.e., 10^{-2} , in order to change from cm to m)

Thus, $\mathbf{k} \approx \begin{bmatrix} 0.39 \\ -16.0 \end{bmatrix} \times 10^6$ and the quadratic spring force-deformation relationship that optimally fits experimental data in the least squares sense is

$$F \approx 38.6 \times 10^4 x - 16.0 \times 10^6 x^2.$$

The data and solution may be plotted with the following commands

```
>> plot(x,f,'o'), hold on
                               % plot data points
>> X = (0:.01:1)*max(x);
>> plot(X,[X' (X.^2)']*k,'-') % best fit curve
>> xlabel('x[m]'), ylabel('F[N]')
```

and the results are shown in Figure 5.

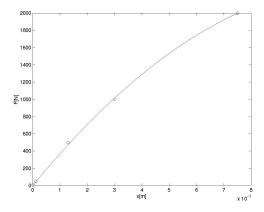


Figure 5: Data for Example 17.1 (circles) and best least squares fit by a quadratic model (solid line).

Matlab has a routine polyfit for data fitting by polynomials: see "help polyfit". It is not applicable in this example because we require that the force - deformation law passes through the origin (so there is no constant term in the quadratic model that we used).

18 Characters, Strings and Text

The ability to process text in numerical processing is useful for the input and output of data to the screen or to disk-files. In order to manage text, a new datatype of "character" is introduced. A piece of text is then simply a string (vector) or array of characters.

Example 18.1 The assignment,

assigns the value A to the 1-by-1 character array t1. The assignment,

```
>> t2 = 'BCDE'
```

assigns the value BCDE to the 1-by-4 character array t2.

Strings can be combined by using the operations for array manipulations.

The assignment,

```
>> t3 = [t1,t2]
```

assigns a value ABCDE to the 1-by-5 character array t3. The assignment,

```
>> t4 = [t3,' are the first 5
'characters in the alphabet.']
```

assigns the value

```
'ABCDE are the first 5 '
```

'characters in the alphabet.'

to the 2-by-27 character array t4. It is essential that the number of characters in both rows of the array t4 is the same, otherwise an error will result. The three dots ... signify that the command is continued on the following line

Sometimes it is necessary to convert a character to the corresponding number, or vice versa. These conversions are accomplished by the commands 'str2num'—which converts a string to the corresponding number, and two functions, 'int2str' and 'num2str', which convert, respectively, an integer and a real number to the corresponding character string. These commands are useful for producing titles and strings, such as 'The value of pi is 3.1416'. This can be generated by the command ['The value of pi is ',num2str(pi)].

```
>> N = 5; h = 1/N;
>> ['The value of N is ',int2str(N),...
', h = ',num2str(h)]
ans =
The value of N is 5, h = 0.2
```

19 Loops

There are occasions that we want to repeat a segment of code a number of different times (such occasions are less frequent than other programming languages because of the : notation).

Example 19.1 Draw graphs of $\sin(n\pi x)$ on the interval $-1 \le x \le 1$ for n = 1, 2, ..., 8.

We could do this by giving 8 separate plot commands but it is much easier to use a loop. The simplest form would be

```
>> x = -1:.05:1;
>> for n = 1:8
     subplot(4,2,n), plot(x,sin(n*pi*x))
and
```

All the commands between the lines starting "for" and "end" are repeated with n being given the value 1 the first time through, 2 the second time, and so forth, until n=8. The subplot constructs a 4×2 array of subwindows and, on the nth time through the loop, a picture is drawn in the nth subwindow.

The commands

```
>> x = -1:.05:1;
>> for n = 1:2:8
     subplot(4,2,n), plot(x,sin(n*pi*x))
     subplot(4,2,n+1), plot(x,cos(n*pi*x))
end
```

draw $\sin n\pi x$ and $\cos n\pi x$ for n=1,3,5,7 alongside each other.

We may use any legal variable name as the "loop counter" (n in the above examples) and it can be made to run through all of the values in a given vector (1:8 and 1:2:8 in the examples).

We may also use for loops of the type

```
>> for counter = [23 11 19 5.4 6] ......
end
```

which repeats the code as far as the end with counter=23 the first time, counter=11 the second time, and so forth.

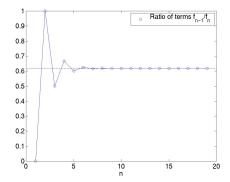
Example 19.2 The Fibonnaci sequence starts off with the numbers 0 and 1, then succeeding terms are the sum of its two immediate predecessors. Mathematically, $f_1 = 0$, $f_2 = 1$ and

$$f_n = f_{n-1} + f_{n-2}, \qquad n = 3, 4, 5, \dots$$

Test the assertion that the ratio f_{n-1}/f_n of two successive values approaches the golden ratio $(\sqrt{5}-1)/2 = 0.6180...$

```
>> F(1) = 0; F(2) = 1;
>> for i = 3:20
    F(i) = F(i-1) + F(i-2);
end
>> plot(1:19, F(1:19)./F(2:20),'o')
>> hold on, xlabel('n')
>> plot(1:19, F(1:19)./F(2:20),'-')
>> legend('Ratio of terms f_{n-1}/f_n')
>> plot([0 20], (sqrt(5)-1)/2*[1,1],'--')
```

The last of these commands produces the dashed horizontal line.



Example 19.3 Produce a list of the values of the sums

$$S_{20} = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \dots + \frac{1}{20^2}$$

$$S_{21} = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \dots + \frac{1}{20^2} + \frac{1}{21^2}$$

$$\vdots$$

$$S_{100} = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \dots + \frac{1}{20^2} + \frac{1}{21^2} + \dots + \frac{1}{100^2}$$

There are a total of 81 sums. The first can be computed using sum(1./(1:20).^2) (The function sum with a vector argument sums its components. See §22.2].) A suitable piece of Matlab code might be

```
>> S = zeros(100,1);
>> S(20) = sum(1./(1:20).^2);
>> for n = 21:100
>> S(n) = S(n-1) + 1/n^2;
>> end
>> clf; plot(S,'.',[20 100],[1,1]*pi^2/6,'-')
>> axis([20 100 1.5 1.7])
>> [ (98:100)' S(98:100)]
ans =
    98.0000    1.6364
    99.0000    1.6365
    100.0000    1.6366
```

where a column vector **S** was created to hold the answers. The first sum was computed directly using the sum command then each succeeding sum was found by adding $1/n^2$ to its predecessor. The little table at the end shows the values of the last three sums—it appears that they are approaching a limit (the value of the limit is $\pi^2/6 = 1.64493...$).

Exercise 19.1 Repeat Example 19.3 to include 181 sums (i.e., the final sum should include the term $1/200^2$.)

20 Logicals

Matlab represents true and false by means of the integers 0 and 1.

If at some point in a calculation a scalar x, say, has been assigned a value, we may make certain logical tests on it:

```
x == 2     is x equal to 2?
x ~= 2     is x not equal to 2?
x > 2     is x greater than 2?
x < 2     is x less than 2?
x >= 2     is x greater than or equal to 2?
```

 $x \le 2$ is x less than or equal to 2? Pay particular attention to the fact that the test for equality involves two equal signs ==.

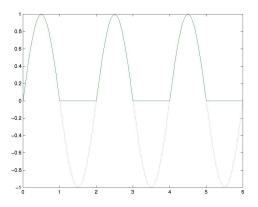
When \mathbf{x} is a vector or a matrix, these tests are performed elementwise:

We may combine logical tests, as in

As one might expect, & represents and and (not so clearly) the vertical bar | means or; also ~ means not as in ~= (not equal), ~(x>0), etc.

One of the uses of logical tests is to "mask out" certain elements of a matrix.

so the matrix pos contains just those elements of x that Method 1: are non-negative.



20.1 While Loops

There are some occasions when we want to repeat a section of Matlab code until some logical condition is satisfied, but we cannot tell in advance how many times we have to go around the loop. This we can do with a while...end construct.

Example 20.1 What is the greatest value of n that can be used in the sum

$$1^2 + 2^2 + \dots + n^2$$

and get a value of less than 100?

The lines of code between while and end will only be executed if the condition $S+ (n+1)^2 < 100$ is true.

Exercise 20.1 Replace 100 in the previous example by 10 and work through the lines of code by hand. You should get the answers n = 2 and S = 5.

Exercise 20.2 Type the code from Example 20.1 into a script-file named WhileSum.m (See §13.)

A more typical example is

Example 20.2 Find the approximate value of the root of the equation $x = \cos x$. (See Example 10.1.)

We may do this by making a guess $x_1 = \pi/4$, say, then computing the sequence of values

$$x_n = \cos x_{n-1}, \qquad n = 2, 3, 4, \dots$$

and continuing until the difference between two successive values $|x_n - x_{n-1}|$ is small enough.

```
>> x = zeros(1,20); x(1) = pi/4;
>> n = 1; d = 1;
>> while d > 0.001
    n = n+1; x(n) = cos(x(n-1));
     d = abs(x(n) - x(n-1));
   end
   n,x
n =
    14
x =
Columns 1 through 7
0.7854 0.7071 0.7602 0.7247 0.7487 0.7326 0.7435
Columns 8 through 14
0.7361 0.7411 0.7377 0.7400 0.7385 0.7395 0.7388
Columns 15 through 20
        0
             0
                  0
                            0
```

There are a number of deficiencies with this program. The vector ${\bf x}$ stores the results of each iteration but we don't know in advance how many there may be. In any event, we are rarely interested in the intermediate values of x, only the last one. Another problem is that we may never satisfy the condition $d \leq 0.001$, in which case the program will run forever—we should place a limit on the maximum number of iterations. Incorporating these improvements leads to

Method 2:

```
>> xold = pi/4; n = 1; d = 1;
>> while d > 0.001 \& n < 20
    n = n+1; xnew = cos(xold);
     d = abs( xnew - xold );
    xold = xnew;
   end
>> [n, xnew, d]
ans =
              0.7388
                        0.0007
   14.0000
```

We continue around the loop so long as d > 0.001 and n < 20. For greater precision we could use the condition d > 0.0001, and this gives

```
>> [n, xnew, d]
   19.0000
               0.7391
                          0.0001
```

from which we may judge that the root required is x =0.739 to 3 decimal places.

The general form of while statement is

```
while a logical test
                  Commands to be executed
                  when the condition is true
end
```

20.2 if...then...else...end

This allows us to execute different commands depending on the truth or falsity of some logical tests. To test whether or not π^e is greater than, or equal to, e^{π} :

```
>> a = pi^exp(1); c = exp(pi);
>> if a >= c
    b = sqrt(a^2 - c^2)
end
```

so that b is assigned a value only if $a \geq c$. There is no output so we deduce that $a=\pi^e < c=e^\pi$. A more common situation is

which ensures that ${\tt b}$ is always assigned a value and confirming that a < c.

A more extended form is

```
>> if a >= c
    b = sqrt(a^2 - c^2)
    elseif a^c > c^a
    b = c^a/a^c
    else
    b = a^c/c^a
    end
b =
    0.2347
```

Exercise 20.3 Which of the above statements assigned a value to b?

The general form of the if statement is

```
if logical test 1
Commands to be executed if test 1 is
true
elseif logical test 2
Commands to be executed if test 2 is
true but test 1 is false
:
end
```

21 Function m-files

These are a combination of the ideas of script m-files (§7) and mathematical functions.

Example 21.1 The area, A, of a triangle with sides of length a, b and c is given by

$$A = \sqrt{s(s-a)(s-b)(s-c)},$$

where s = (a + b + c)/2. Write a Matlab function that will accept the values a, b and c as inputs and return the value of A as output.

The main steps to follow when defining a Matlab function are:

- Decide on a name for the function, making sure that it does not conflict with a name that is already used by Matlab. In this example the name of the function is to be area, so its definition will be saved in a file called area.m
- 2. The first line of the file must have the format: function [list of outputs]

 $= function_name(list of inputs)$

For our example, the output (A) is a function of the three variables (inputs) a, b and c so the first line should read

```
function [A] = area(a,b,c)
```

- 3. Document the function. That is, describe briefly the purpose of the function and how it can be used. These lines should be preceded by % which signify that they are comment lines that will be ignored when the function is evaluated.
- Finally include the code that defines the function. This should be interspersed with sufficient comments to enable another user to understand the processes involved.

The complete file might look like:

```
function [A] = area(a,b,c)
% Compute the area of a triangle whose
% sides have length a, b and c.
% Inputs:
    a,b,c: Lengths of sides
% Output:
%
        A: area of triangle
% Usage:
%
      Area = area(2,3,4);
% Written by dfg, Oct 14, 1996.
s = (a+b+c)/2;
A = sqrt(s*(s-a)*(s-b)*(s-c));
The command
>> help area
will produce the leading comments from the file:
  Compute the area of a triangle whose
  sides have length a, b and c.
  Inputs:
    a,b,c: Lengths of sides
  Output:
         A: area of triangle
  Usage:
      Area = area(2,3,4);
  Written by dfg, Oct 14, 1996.
```

To evaluate the area of a triangle with side of length 10, 15, 20:

```
>> Area = area(10,15,20)
Area =
72.6184
```

where the result of the computation is assigned to the variable Area. The variable s used in the definition of the function above is a "local variable": its value is local to the function and cannot be used outside:

>> s

??? Undefined function or variable s.

If we were to be interested in the value of s as well as A, then the first line of the file should be changed to

function
$$[A,s] = area(a,b,c)$$

where there are two output variables.

This function can be called in several different ways:

1. No outputs assigned

```
>> area(10,15,20)
ans =
72.6184
```

gives only the area (first of the output variables from the file) assigned to ans; the second output is ignored.

2. One output assigned

```
>> Area = area(10,15,20)
Area =
72.6184
```

again the second output is ignored.

3. Two outputs assigned

```
>> [Area, hlen] = area(10,15,20)
Area =
    72.6184
hlen =
    22.5000
```

Exercise 21.1 In any triangle the sum of the lengths of any two sides cannot exceed the length of the third side. The function area does not check to see if this condition is fulfilled (try area(1,2,4)). Modify the file so that it computes the area only if the sides satisfy this condition.

21.1 Examples of functions

We revisit the problem of computing the Fibonnaci sequence defined by $f_1 = 0, f_2 = 1$ and

$$f_n = f_{n-1} + f_{n-2}, \qquad n = 3, 4, 5, \dots$$

We want to construct a function that will return the nth number in the Fibonnaci sequence f_n .

• Input: Integer n

• Output: f_n

We shall describe four possible functions and try to assess which provides the best solution.

Method 1: File Fib1.m

```
function f = Fib1(n)
% Returns the nth number in the
% Fibonacci sequence.
F=zeros(1,n+1);
F(2) = 1;
  for i = 3:n+1
    F(i) = F(i-1) + F(i-2);
  end
f = F(n);
```

This code resembles that given in Example 19.2. We have simply enclosed it in a function m-file and given it the appropriate header,

Method 2: File Fib2.m

The first version was rather wasteful of memory—it saved all the entries in the sequence even though we only required the last one for output. The second version removes the need to use a vector.

```
function f = Fib2(n)
% Returns the nth number in the
% Fibonacci sequence.
if n==1
    f = 0;
elseif n==2
    f = 1;
else
    f1 = 0; f2 = 1;
    for i = 2:n-1
        f = f1 + f2;
        f1=f2; f2 = f;
    end
end
```

Method 3: File: Fib3.m

This version makes use of an idea called "recursive programming"—the function makes calls to itself.

```
function f = Fib3(n)
% Returns the nth number in the
% Fibonacci sequence.
if n==1
    f = 0;
elseif n==2
    f = 1;
else
    f = Fib3(n-1) + Fib3(n-2);
end
```

Method 4: File Fib4.m

The final version uses matrix powers. The vector \mathbf{y} has two components, $\underline{y} = \begin{bmatrix} f_n \\ f_{n+1} \end{bmatrix}$.

```
function f = Fib4(n)
% Returns the nth number in the
% Fibonacci sequence.
A = [0 1;1 1];
y = A^n*[1;0];
f=y(1);
```

Assessment: One may think that, on grounds of style, the 3rd is best (it avoids the use of loops) followed by the second (it avoids the use of a vector). The situation is much different when it cames to speed of execution. When n=20 the time taken by each of the methods is (in seconds)

Method	Time
1	0.0118
2	0.0157
3	36.5937
4	0.0078

It is impractical to use Method 3 for any value of n much larger than 10 since the time taken by method 3 almost doubles whenever n is increased by just 1. When n=150

Method	Time
1	0.0540
2	0.0891
3	_
4	0.0106

Clearly the 4th method is much the fastest.

22 Further Built-in Functions

22.1 Rounding Numbers

There are a variety of ways of rounding and chopping real numbers to give integers. Use the definitions given in the table in §28 on page 32 in order to understand the output given below:

Do "help round" for help information.

22.2 The sum Function

The "sum" applied to a vector adds up its components (as in sum(1:10)) while, for a matrix, it adds up the components in **each column** and returns a row vector. sum(sum(A)) then sums all the entries of A.

The sums of squares of the entries in each column of A are equal to 1 and the sum of squares of all the entries is equal to 3.

It appears that the products AA' and A'A are both equal to the identity:

This is confirmed since the differences are at round—off error levels (less than 10^{-15}). A matrix with this property is called an *orthogonal* matrix.

22.3 max & min

These functions act in a similar way to sum. If x is a vector, then max(x) returns the largest element in x

```
2.4000
>> [m, j] = max(x)
m =
2.3000
j =
```

When we ask for two outputs, the first gives us the maximum entry and the second the index of the maximum element.

For a matrix, A, max(A) returns a row vector containing the maximum element from each column. Thus to find the largest element in A we have to use max(max(A)).

22.4 Random Numbers

The function rand(m,n) produces an $m \times n$ matrix of random numbers, each of which is in the range 0 to 1. rand on its own produces a single random number.

```
>> y = rand, Y = rand(2,3)
y =
0.9191
Y =
0.6262 0.1575 0.2520
0.7446 0.7764 0.6121
```

Repeating these commands will lead to different answers.

Example: Write a function-file that will simulate n throws of a pair of dice.

This requires random numbers that are integers in the range 1 to 6. Multiplying each random number by 6 will give a real number in the range 0 to 6; rounding these to whole numbers will not be correct since it will then be possible to get 0 as an answer. We need to use

```
floor(1 + 6*rand)
```

Recall that floor takes the largest integer that is smaller than a given real number (see Table 2, page 32).

File: dice.m

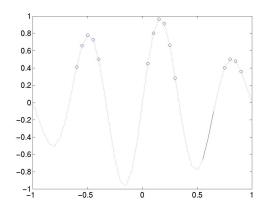
```
function [d] = dice(n)
% simulates "n" throws of a pair of dice
% Input:
             n, the number of throws
% Output:
             an n times 2 matrix, each row
%
             referring to one throw.
%
% Useage: T = dice(3)
   d = floor(1 + 6*rand(n,2));
%% end of dice
>> dice(3)
ans =
     6
           1
     2
           3
     4
           1
>> sum(dice(100))/100
ans =
   3.8500
              3.4300
```

The last command gives the average value over 100 throws (it should have the value 3.5).

22.5 find for vectors

The function "find" returns a list of the positions (indices) of the elements of a vector satisfying a given condition. For example,

```
>> x = -1:.05:1;
>> y = \sin(3*pi*x).*exp(-x.^2); plot(x,y,':')
>> k = find(y > 0.2)
 k =
  Columns 1 through 12
   9 10 11 12 13 22
                        23 24
                                 25 26 27 36
  Columns 13 through 15
  37 38 39
>> hold on, plot(x(k),y(k),'o')
>> km = find(x>0.5 \& y<0)
km
    32
               34
          33
>> plot(x(km),y(km),'-')
```



22.6 find for matrices

The find-function operates in much the same way for matrices:

```
>> A = [ -2 3 4 4; 0 5 -1 6; 6 8 0 1]
A =

-2 3 4 4
0 5 -1 6
6 8 0 1
>> k = find(A==0)
k =

2
```

Thus, we find that A has elements equal to 0 in positions 2 and 9. To interpret this result we have to recognize that "find" first reshapes A into a column vector—this is equivalent to numbering the elements of A by columns as in

```
>> n = find(A <= 0)
n =
```

1	
2	
8	
9	
>> A(n)	
ans =	
-2	
0	
-1	
0	

Thus, n gives a list of the locations of the entries in A that are ≤ 0 and then A(n) gives us the values of the elements selected.

Since we are dealing with A^{γ} , the entries are numbered by rows.

23 Plotting Surfaces

A surface is defined mathematically by a function f(x, y)corresponding to each value of (x, y) we compute the
height of the function by

$$z = f(x, y).$$

In order to plot this we have to decide on the ranges of x and y—suppose $2 \le x \le 4$ and $1 \le y \le 3$. This gives us a square in the (x,y)-plane. Next, we need to choose a grid on this domain; Figure 6 shows the grid with intervals 0.5 in each direction. Finally, we have

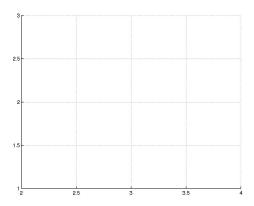


Figure 6: An example of a 2D grid

to evaluate the function at each point of the grid and "plot" it.

Suppose we choose a grid with intervals 0.5 in each direction for illustration. The x- and y-coordinates of the grid lines are

$$x = 2:0.5:4; y = 1:0.5:3;$$

in Matlab notation. We construct the grid with meshgrid: function.

If we think of the ith point along from the left and the jth point up from the bottom of the grid) as corresponding to the (i,j)th entry in a matrix, then (X(i,j), Y(i,j)) are the coordinates of the point. We then need to evaluate the function f using X and Y in place of x and y, respectively.

Example 23.1 Plot the surface defined by the function

$$f(x,y) = (x-3)^2 - (y-2)^2$$

for $2 \le x \le 4$ and $1 \le y \le 3$.

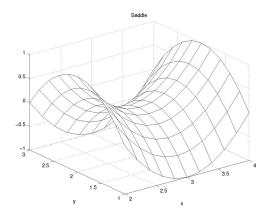


Figure 7: Plot of Saddle function.

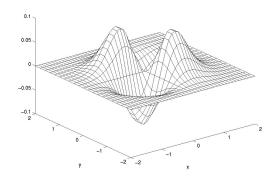
Exercise 23.1 Repeat the previous example replacing mesh by surf and then by surfl. Consult the help pages to find out more about these functions.

 ${\bf Example~23.2~} \textit{Plot the surface defined by the function} \\$

$$f = -xye^{-2(x^2 + y^2)}$$

on the domain $-2 \le x \le 2, -2 \le y \le 2$. Find the values and locations of the maxima and minima of the

```
>> [X,Y] = meshgrid(-2:.1:2,-2:.2:2);
>> f = -X.*Y.*exp(-2*(X.^2+Y.^2));
>> figure (1)
>> mesh(X,Y,f), xlabel('x'), ylabel('y'), grid
>> figure (2), contour(X,Y,f)
>> xlabel('x'), ylabel('y'), grid, hold on
```



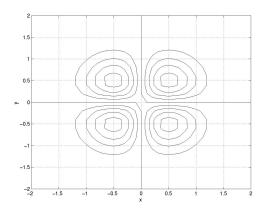


Figure 8: "mesh" and "contour" plots.

To locate the maxima of the " \mathbf{f} " values on the grid:

```
>> fmax = max(max(f))
fmax =
          0.0886
>> kmax = find(f==fmax)
kmax =
          323
          539
>> Pos = [X(kmax), Y(kmax)]
Pos =
          -0.5000     0.6000
          0.5000     -0.6000
>> plot(X(kmax),Y(kmax),'*')
>> text(X(kmax),Y(kmax),' Maximum')
```

24 Timing

Matlab allows the timing of sections of code by providing the functions tic and toc. tic switches on a stopwatch while toc stops it and returns the CPU time

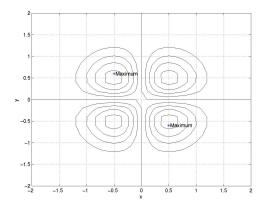


Figure 9: contour plot showing maxima.

(Central Processor Unit) in seconds. The timings will vary depending on the model of computer being used and its current load.

```
\Rightarrow tic,for j=1:1000,x = pi*R(3);end,toc
elapsed_time =
                   0.5110
>> tic,for j=1:1000,x=pi*R(3);end,toc
elapsed_time =
                  0.5017
>> tic,for j=1:1000,x=R(3)/pi;end,toc
elapsed_time =
                  0.5203
>> tic,for j=1:1000,x=pi+R(3);end,toc
elapsed_time =
                  0.5221
>> tic,for j=1:1000,x=pi-R(3);end,toc
elapsed_time =
                  0.5154
>> tic,for j=1:1000,x=pi^R(3);end,toc
elapsed_time =
                  0.6236
```

25 On-line Documentation

In addition to the on–line help facility, there is a hypertext browsing system giving details of (most) commands and some examples. This is accessed by

>> doc

which brings up the *Netscape* document previewer (and allows for "surfing the internet superhighway"—the World Wide Web (WWW). It is connected to a worldwide system which, given the appropriate addresses, will provide information on almost any topic).

Words that are underlined in the browser may be clicked on with LB and lead to either a further subindex or a help page.

Scroll down the page shown and click on <u>general</u> which will take you to "General Purpose Commands"; click on <u>clear</u>. This will describe how you can clear a variable's value from memory.

You may then either click the "Table of Contents" which takes you back to the start, "Index" or the Back button at the lower left corner of the window which will take you back to the previous screen.

To access other "home pages", click on Open at the bottom of the window and, in the "box" that will open up, type

http://www.maths.dundee.ac.uk

or

http://www.maths.dundee.ac.uk/software/

and select Matlab from the array of choices.

26 Reading and Writing Data **Files**

Direct input of data from keyboard becomes impractical when

- the amount of data is large and
- the same data is analysed repeatedly.

In these situations input and output is preferably accomplished via data files. We have already described in Section 9 the use of the commands save and load that, respectively, write and read the values of variables to disk files.

When data are written to or read from a file it is crucially important that a correct data format is used. The data format is the key to interpreting the contents of a file and must be known in order to correctly interpret the data in an input file. There a two types of data files: formatted and unformatted. Formatted data files uses format strings to define exactly how and in what positions of a record the data is stored. Unformatted storage, on the other hand, only specifies the number format.

The files used in this section are available from the web

http://www.maths.dundee.ac.uk/software/#matlab

Those that are unformatted are in a satisfactory form for the Windows version on Matlab (version 6.1) but Step 1: Assign a namestring to a file identifier. not on Version 5.3 under Unix.

Exercise 26.1 Suppose the numeric data is stored in a file 'table.dat' in the form of a table, as shown below.

The three commands,

```
>> fid = fopen('table.dat','r');
>> a = fscanf(fid,'%3d%4d');
>> fclose(fid);
```

respectively

1. open a file for reading (this is designated by the string 'r'). The variable fid is assigned a unique integer which identifies the file used (a file identifier). We use this number in all subsequent references to the file.

- 2. read pairs of numbers from the file with file identifier fid, one with 3 digits and one with 4 digits,
- 3. close the file with file identifier fid.

This produces a **column vector** a with elements, 100 2256 200 4564 ...500 6432. This vector can be converted to 5×2 matrix by the command A = reshape(2,2,5);

Formatted Files 26.1

Some computer codes and measurement instruments produce results in formatted data files. In order to read these results into Matlab for further analysis the data format of the files must be known. Formatted files in ASCII format are written to and read from with the commands fprintf and fscanf.

> fprintf(fid, 'format', variables) writes variables an a format specified in string 'format' to the file with identifier fid

> a = fscanf(fid, 'format', size) assigns to variable a data read from file with identifier fid under format 'format'.

Exercise 26.2 Study the available information and help on fscanf and fprintf commands. What is the meaning of the format string, '%3d\n'?

Example 26.1 Suppose a sound pressure measurement system produces a record with 512 time - pressure readings stored on a file 'sound.dat'. Each reading is listed on a separate line according to a data format specified by the string, '%8.6f %8.6f'.

A set of commands reading time – sound pressure data from 'sound.dat' is,

```
>> fid1 = fopen('sound.dat', 'r');
```

The string 'r' indicates that data is to be read (not written) from the file.

Step 2: Read the data to a vector named 'data' and close the file.

```
>> data = fscanf(fid1, '%f %f');
>> fclose(fid1);
```

Step 3: Partition the data in separate time and sound pressure vectors,

```
>> t = data(1:2:length(data));
>> press = data(2:2:length(data));
```

The pressure signal can be plotted in a lin-lin diagram,

```
>> plot(t, press);
```

The result is shown in Figure 10.

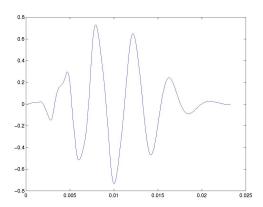


Figure 10: Graph of "sound data" from Example 26.1

26.2 Unformatted Files

Unformatted or binary data files are used when small-sized files are required. In order to interpret an unformatted data file the data precision must be specified. The precision is specified as a string, e.g., 'float32', controlling the number of bits read for each value and the interpretation of those bits as character, integer or floating point values. Precision 'float32', for instance, specifies each value in the data to be stored as a floating point number in 32 memory bits.

Example 26.2 Suppose a system for vibration measurement stores measured acceleration values as floating point numbers using 32 memory bits. The data is stored on file 'vib.dat. The following commands illustrate how the data may be read into Matlab for analysis.

Step 1: Assign a file identifier, fid, to the string specifying the file name.

```
>> fid = fopen('vib.dat', 'rb');
```

The string 'rb' specifies that binary numbers are to be read from the file.

Step 2 Read all data stored on file 'vib.dat' into a vector vib.

The size(vib) command determines the size, i.e., the number of rows and columns of the vibration data vector.

In order to plot the vibration signal with a correct time scale, the sampling frequency (the number of instrument readings taken per second) used by the measurement system must be known. In this case it is known to be 24000 Hz so that there is a time interval of 1/24000 seconds between two samples.

Step 3: Create a column vector containing the correct time scale.

```
>> dt = 1/24000;
>> t = dt*(1:length(vib))';
```

Step 4: Plot the vibration signal in a lin-lin diagram

```
>> plot(t,vib);
>> title('Vibration signal');
>> xlabel('Time,[s]');
>> ylabel('Acceleration, [m/s^2]');
```

27 Graphic User Interfaces

The efficiency of programs that are used often and by several different people can be improved by simplifying the input and output data management. The use of Graphic User Interfaces (GUI), which provides facilities such as menus, pushbuttons, sliders etc, allow programs to be used without any knowledge of Matlab. They also provides means for efficient data management.

A graphic user interface is a Matlab script file customized for repeated analysis of a specific type of problem. There are two ways to design a graphic user interface. The simplest method is to use a tool especially designed for the purpose. Matlab provides such a tool and it is invoked by typing 'guide' at the Matlab prompt. Maximum flexibility and control over the programming is, however, obtained by using the basic user interface commands. The following text demonstrates the use of some basic commands.

Example 27.1 Suppose a sound pressure spectrum is to be plotted in a graph. There are four alternative plot formats; lin-lin, lin-log, log-lin and log-log. The graphic user interface below reads the pressure data stored on a binary file selected by the user, plots it in a lin-lin format as a function of frequency and lets the user switch between the four plot formats.

We use two m-files. The first (specplot.m) is the main driver file which builds the graphics window. It calls the second file (firstplot.m) which allows the user to select among the possible *.bin files in the current directory.

```
% Create pushbuttons for switching between four Executing this GUI from the command line
% different plot formats. Set up the axis stings.(>> specplot) brings the following screen.
X = 'Frequency, [Hz]';
Y = 'Pressure amplitude, [Pa]';
linlinBtn = uicontrol('style', 'pushbutton',...
  'string','lin-lin',...
  'position', [200,395,40,20], 'callback',...
  'plot(fdat,pdat);xlabel(X);ylabel(Y);');
linlogBtn = uicontrol('style','pushbutton',...
  'string','lin-log',...
  'position',[240,395,40,20],...
  'callback',...
  'semilogy(fdat,pdat);xlabel(X);ylabel(Y);');
loglinBtn = uicontrol('style', 'pushbutton',...
  'string','log-lin',...
  'position',[280,395,40,20],...
  'callback',...
  'semilogx(fdat,pdat);xlabel(X);ylabel(Y);');
loglogBtn = uicontrol('style', 'pushbutton',...
  'string','log-log',...
  'position',[320,395,40,20],...
  'callback',...
                                   ylabel(Y);');
  'loglog(fdat,pdat);xlabel(X);
% Create exit pushbutton with red text.
exitBtn = uicontrol('Style', 'pushbutton',...
  'string','EXIT','position',[510,395,40,20],...
  'foregroundcolor',[1 0 0],'callback','close;')
% Script file: firstplot.m
% Brings template for file selection. Reads
\% selected filename and path and plots
% spectrum in a lin-lin diagram.
% Output data are frequency and pressure
% amplitude vectors: 'fdat' and 'pdat'.
% Author: U Carlsson, 2001-08-22
function [fdat,pdat] = firstplot
% Call Matlab function 'uigetfile' that
% brings file selction template.
[filename,pathname] = uigetfile('*.bin',...
  'Select binary data-file:');
% Change directory
cd(pathname);
% Open file for reading binary floating
% point numbers.
fid = fopen(filename, 'rb');
data = fread(fid, 'float32');
% Close file
fclose(fid);
% Partition data vector in frequency and
% pressure vectors
pdat = data(2:2:length(data));
fdat = data(1:2:length(data));
% Plot pressure signal in a lin-lin diagram
plot(fdat,pdat);
% Define suitable axis labels
xlabel('Frequency, [Hz]');
ylabel('Pressure amplitude, [Pa]');
```

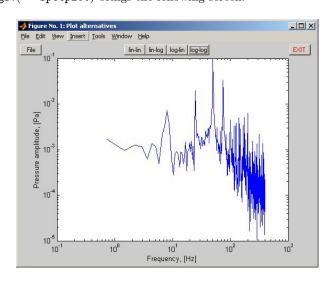


Figure 11: Graph of "vibration data" from Example 27.1

Example 27.1 illustrates how the 'callback' property allows the programmer to define what actions should result when buttons are pushed etc. These actions may consist of single Matlab commands or complicated sequences of operations defined in various subroutines.

Exercise 27.1 Five different sound recordings are stored on binary data files, sound1.bin, sound2.bin, ..., sound5.bin. The storage precision is 'float32' and the sounds are recorded with sample frequency 12000 Hz. Write a graphic user interface that, opens an interface

- lets the user select one of the five sounds,
- plots the selected sound pressure signal as a function of time in a lin-lin diagram,
- lets the user listen to the sound by pushing a 'SOUND' button and finally
- closes the session by pressing a 'CLOSE' button.

28 Command Summary

The command

window and

>> help

will give a list of categories for which help is available (e.g. matlab/general covers the topics listed in Table 3. Further information regarding the commands listed in this section may then be obtained by using:

>> help topic try, for example, >> help help

'n		
	abs	Absolute value
	sqrt	Square root function
ı	sign	Signum function
ı	conj	Conjugate of a complex number
ı	imag	Imaginary part of a complex
ı		number
ı	real	Real part of a complex number
ı	angle	Phase angle of a complex number
	cos	Cosine function
ı	sin	Sine function
	tan	Tangent function
ı	exp	Exponential function
ı	log	Natural logarithm
ı	log10	Logarithm base 10
	cosh	Hyperbolic cosine function
	sinh	Hyperbolic sine function
ı	tanh	Hyperbolic tangent function
ı	acos	Inverse cosine
ı	acosh	Inverse hyperbolic cosine
ı	asin	Inverse sine
ı	asinh	Inverse hyperbolic sine
ı	atan	Inverse tan
ı	atan2	Two-argument form of inverse
		tan
	atanh	Inverse hyperbolic tan
ı	round	Round to nearest integer
ı	floor	Round towards minus infinity
	fix	Round towards zero
	ceil	Round towards plus infinity
	rem	Remainder after division

Table 2: Elementary Functions

Mana	aging commands and functions.	
help	On-line documentation.	
doc	Load hypertext documentation.	
what	Directory listing of M-, MAT-	
wnat	and MEX-files.	
type	List M-file.	
lookfor	v G	
	HELP entries.	
which	Locate functions and files.	
demo	Run demos.	
Manag	ing variables and the workspace.	
who	List current variables.	
whos	List current variables, long form.	
load	Retrieve variables from disk.	
save	Save workspace variables to disk.	
clear	Clear variables and functions	
	from memory.	
size	Size of matrix.	
length	Length of vector.	
disp	Display matrix or text.	
Working with files and the operating system.		
cd	Change current working direc-	
	tory.	
dir	Directory listing.	
delete	Delete file.	
!	Execute operating system com-	
	mand.	
unix	Execute operating system com-	
	mand & return result.	
diary	Save text of MATLAB session.	
	trolling the command window.	
cedit	Set command line edit/recall fa-	
	cility parameters.	
clc	Clear command window.	
home	Send cursor home.	
format	Set output format.	
echo	Echo commands inside script	
30113	files.	
more	Control paged output in com-	
mor o	mand window.	
	Quitting from MATLAB.	
quit	Terminate MATLAB.	
quit	reminate matlad.	

Table 3: General purpose commands.

	26	
	Matrix analysis.	
cond	Matrix condition number.	
norm	Matrix or vector norm.	
rcond	LINPACK reciprocal condition	
	estimator.	
rank	Number of linearly independent	
	rows or columns.	
det	Determinant.	
trace	Sum of diagonal elements.	
null	Null space.	
orth	Orthogonalization.	
rref	Reduced row echelon form.	
	Linear equations.	
\ and /	Linear equation solution; use	
	"help slash".	
chol	Cholesky factorization.	
lu	Factors from Gaussian elimina-	
	tion.	
inv	Matrix inverse.	
qr	Orthogonal- triangular decompo-	
	sition.	
qrdelete	Delete a column from the QR fac-	
_	torization.	
qrinsert	Insert a column in the QR factor-	
_	ization.	
nnls	Non-negative least- squares.	
pinv	Pseudoinverse.	
lscov	Least squares in the presence of	
	known covariance.	
Eigenvalues and singular values.		
eig	Eigenvalues and eigenvectors.	
poly	Characteristic polynomial.	
polyeig	Polynomial eigenvalue problem.	
hess	Hessenberg form.	
qz	Generalized eigenvalues.	
rsf2csf	Real block diagonal form to com-	
	plex diagonal form.	
cdf2rdf	Complex diagonal form to real	
	block diagonal form.	
schur	Schur decomposition.	
balance	Diagonal scaling to improve	
	eigenvalue accuracy.	
svd	Singular value decomposition.	
	Matrix functions.	
expm	Matrix exponential.	
expm1	M- file implementation of expm.	
expm2	Matrix exponential via Taylor se-	
1 1	ries.	
expm3	Matrix exponential via eigenval-	
22.70	ues and eigenvectors.	
logm	Matrix logarithm.	
sqrtm	Matrix square root.	
funm	_	
ı uiill	Evaluate general matrix function.	

 $\label{thm:constant} \begin{tabular}{ll} Table 4: Matrix functions—numerical linear algebra. \end{tabular}$

	Crapping & plotting		
	Graphics & plotting.		
figure	Create Figure (graph window).		
clf	Clear current figure.		
close	Close figure.		
subplot	Create axes in tiled positions.		
axis	Control axis scaling and appear-		
	ance.		
hold	Hold current graph.		
figure	Create figure window.		
text	Create text.		
print	Save graph to file.		
plot	Linear plot.		
loglog	Log-log scale plot.		
semilogx	Semi-log scale plot.		
semilogy	Semi-log scale plot.		
Specialized X-Y graphs.			
polar	Polar coordinate plot.		
bar	Bar graph.		
stem	Discrete sequence or "stem" plot.		
stairs	Stairstep plot.		
errorbar	Error bar plot.		
hist	Histogram plot.		
rose	Angle histogram plot.		
compass	Compass plot.		
feather	Feather plot.		
fplot	Plot function.		
comet	Comet-like trajectory.		
Graph annotation.			
title	Graph title.		
xlabel	X-axis label.		
ylabel	Y-axis label.		
text	Text annotation.		
gtext	Mouse placement of text.		
grid	Grid lines.		
contour	Contour plot.		
mesh	3-D mesh surface.		
surf	3-D shaded surface.		
waterfall	Waterfall plot.		
view	3-D graph viewpoint specifica-		
	tion.		
zlabel	Z-axis label for 3-D plots.		
gtext	Mouse placement of text.		
grid	Grid lines.		
0	0114 111100.		

Table 5: Graphics & plot commands.

\mathbf{Index}

<, 21, 23	echo, 10
<=, 21, 23 <=, 21, 23	elementary functions, 4
==, 21, 23	eye, 14
	eye, 14
>, 21, 23	folgo 01
>=, 21, 23	false, 21
% , 10, 23	Fibonnaci, 20, 24
', 5	figure, 8
., 6	file
.*, 11	function, 23
, 8	script, 10
./, 12	find, 26
.^, 12	fix, 32
:, 5, 16	floor, 32
	floor, 26
;, 3	
1 00	for
abs, 32	loop, 20
accelerators	format, 3
keyboard, 9	long, 12
and, 21	function m-files, 23
angle, 32	functions
ans, 3	elementary, 4
array, 13	trigonometric, 4
axes, 9, 13	erigenemetrie, i
	get, 8
axis, 9	9 ,
auto, 9	graphs, see plotting
normal, 9	grid, 7, 13, 27
$\mathtt{square},9$	GUI, 30
	–
browser, 28	hard copy, 7
	help, 2, 23
ceil, 32	hold, 7, 13
clf, 8	home page, 28
close, 8	
colon notation, 5, 16	if statement, 22
column vectors, 5	imag, 32
And the second s	G,
comment (%), 10, 23	keyboard accelerators, 9
complex	.,
conjugate transpose, 5	labels for plots, 6
numbers, 5	legend, 7
complex numbers, 3	_
components of a vector, 4	length of a vector, 4, 5, 10
conj, 32	line styles, 7
contour, 27	logical conditions, 21
copying output, 9	loops, 20
$\cos, 32$	while, 22
CPU, 28	
	m-files, 10, 23
cursor keys, 9	matrix, 13
	building, 15
Jama 9	
demo, 3	
diag, 15	diagonal, 14
$\begin{array}{ll} \texttt{diag}, 15 \\ \texttt{diary}, 6 \end{array}$	diagonal, 14 identity, 14
diag, 15	diagonal, 14 identity, 14 indexing, 15
$\begin{array}{ll} \texttt{diag}, 15 \\ \texttt{diary}, 6 \end{array}$	diagonal, 14 identity, 14 indexing, 15 n=5;tridiagonal, 17
$\begin{array}{l} \operatorname{diag}, 15 \\ \operatorname{diary}, 6 \\ \operatorname{dice}, 26 \end{array}$	diagonal, 14 identity, 14 indexing, 15 n=5;tridiagonal, 17 orthogonal, 25
diag, 15 diary, 6 dice, 26 divide dot, 12	diagonal, 14 identity, 14 indexing, 15 n=5;tridiagonal, 17 orthogonal, 25 size, 13
diag, 15 diary, 6 dice, 26 divide dot, 12 documentation, 28	diagonal, 14 identity, 14 indexing, 15 n=5;tridiagonal, 17 orthogonal, 25 size, 13 sparse, 17
diag, 15 diary, 6 dice, 26 divide dot, 12 documentation, 28 dot	diagonal, 14 identity, 14 indexing, 15 n=5;tridiagonal, 17 orthogonal, 25 size, 13
diag, 15 diary, 6 dice, 26 divide dot, 12 documentation, 28 dot divide ./, 12	diagonal, 14 identity, 14 indexing, 15 n=5;tridiagonal, 17 orthogonal, 25 size, 13 sparse, 17
diag, 15 diary, 6 dice, 26 divide dot, 12 documentation, 28 dot divide ./, 12 power .^, 12	diagonal, 14 identity, 14 indexing, 15 n=5;tridiagonal, 17 orthogonal, 25 size, 13 sparse, 17 special, 14
diag, 15 diary, 6 dice, 26 divide dot, 12 documentation, 28 dot divide ./, 12	diagonal, 14 identity, 14 indexing, 15 n=5;tridiagonal, 17 orthogonal, 25 size, 13 sparse, 17 special, 14 spy, 15

1.4	
zeros, 14	strings, 6
matrix products, 17	$\mathtt{subplot},8,20$
matrix-vector products, 16	subscripts, 8
$\max, 25, 27$	sum, 21, 25
$\operatorname{mesh}, 27$	superscripts, 8
meshgrid, 27	surfing the internet highway, 28
$\min, 25, 27$	
more, 2	timing, 28
multi-plots, 7	title for plots, 6
F, ·	toc, 28
Netscape, 28	transposing, 5
norm of a vector, 11	tridiagonal, 17
not, 21, 23	trigonometric functions, 4
numbers, 3	true, 21
complex, 3	type (list contents of m-file), 10
format, 3	
random, 26	variable names, 3
rounding, 25	vector
	components, 4
ones, 14	vectors
or, 21	column, 5
,	row, 4
plot, 20	10,1,1
plotting, 6, 12, 27	what, 10
labels, 6	while loops, 22
line styles, 7	whos, 6
printing, 7	WWW, 28
surfaces, 27	0 0
title, 6	xlabel, 6, 27
power	xterm, 2
dot, 12	
printing plots, 7	ylabel, 6
priorities	
in arithmetic, 3	zeros, 14
product	zoom, 8
dot, 11, 16	
scalar, 16, 17	
0	
quit, 2	
. 00	
rand, 26	
random numbers, 26	
real, 32	
rem, 32	
round, 32	
rounding error, 4	
rounding numbers, 25	
save, 6	
scalar product, 10, 16, 17	
script files, 10	
semi-colon, 4, 13	
set, 9	
shapes, 7	
sign, 32	
\sin , 32	
size, 13	
sort, 5	
sparse, 17	
± /	
spdiags, 17	
spdiags, 17	
spdiags, 17 spy , 15 sqrt, 32	