Institut for Matematik og Datalogi
Syddansk Universitet

November 14, 2001
JFB

Introduction to Information Technology
E01 – Note 6

Lecture, November 2

We continued in chapter 3 of the textbook, concentrating on sections 3.1, 3.2, and 3.4, though using sequential search and binary search as examples. In addition, some types in Maple (for example, integers, reals, sets, lists, lists of lists, and functions) were demonstrated and discussed.

Lecture, November 16

We will cover sections 3.5.2 and 3.6 of chapter 3, plus give an example from computational molecular biology on protein alignment. If there is time, we will begin on section 14.7 of chapter 14. Spreadsheets and statistics in Maple will be demonstrated.

Lecture, November 30

We will continue with chapter 14, with emphasis on section 14.7 and cryptography. Sorting from the algorithms lab will be discussed.

Primary Lab 6 - for week 47

There are several goals for this lab. You should learn:

- how to import data from text files into Maple,
- how to import data from Excel into a Maple spreadsheet,
- more about types in Maple, and
- how to use more of the statistical capabilities in Maple.

You may skip exercises 2 and 7 if you do not have enough time. Do all of the others, including one of the two versions of exercise 8. If you do the second version, looking at the directions for the first may be helpful.

Exercise 1

Importing data and plotting it. Start by opening the Maple 7 program and typing `restart;`. The `restart` command is useful when you want to change your worksheet a little and execute it again; it clears all the variables and assignments. (When you want to execute the entire worksheet again, you can do it through the **Edit** menu button and **Execute**.) Start up the statistical package in Maple by typing `with(stats);`. Read the regression data you used in your last Excel lab into a variable T. Use the function *importdata* to read it from the text file you saved (either from e-mail or from the Web).

Type `T[1,3];` to see the value of the third number in the first list. Try looking at other values, too.

Type `with(statplots);` so that you can easily access those functions. First try plotting the points with `scatterplot(T);`, or add a second argument `color=blue` to change to a darker color which is easier to see. You can change the shape of the symbols marking the points or add a legend by right clicking on the plot.

Exercise 2 – Optional –Do if you have time

To get a histogram of the values in the second list, you can type

```
histogram(T[2],color=yellow);
```

(of course, you can choose a different color or leave it out). The bars all have the same area, but often you may prefer that they have the same width. This can be done with an argument `area=1` or `area=count`. Try both to see what happens. What is the difference? If you want a different number of bars, you can add an argument to specify this.

Exercise 3

Linear regression. In order to perform linear regression on your data, you will need to have the type "list of lists" (which T is not). You can do this with right clicking on the output for T. Assign the result to a variable L. Now type `with(fit);` to make it easier to use the functions there. You can find the equation of the line computed by the least squares method by typing `lin:=fit[leastsquare[[x,y]]](L);` This uses the data in L and gives the variables involved the names x and y. Compare this equation to the one calculated by Excel in that laboratory exercise. It should be about the same. The coefficient computed by Excel is the square of the linear correlation, so you can get that by typing `with(describe);` and then `r:=linearcorrelation(T[1],T[2]);` and `r*r;`. Compare with the result from Excel.

Exercise 4

Try fitting other types of curves (other than linear) to your data. For example, to use the least squares method to solve for the best a, b, and c for a curve of the form $a{\cdot}x{\cdot}ln(x){+}b{\cdot}x{+}c$

to fit the points, type

```
eqn:=fit[leastsquare[[x,y],y=a*x*ln(x) + b*x +c,{a,b,c}]](L);.
```

Try using the least squares method to find the best fit for a curve of the form

$$a \cdot x \cdot (ln(x))^2 + b \cdot x \cdot ln(x) + c \cdot x + d.$$

Exercise 5

Plotting the data without the statistics package. To plot the data in your variable T using the standard plotting function in Maple, you need to create a list of points (pairs of numbers) from your original data. You need to process the two lists you have in T, looking at one element from each list at a time, and creating a pair from them. To do this, you first create a function that takes two elements and makes a pair from them: `pair := (x,y) -> [x,y];`. Try applying the function to two arguments, to see how it works: `pair(t,z);` or `pair(15,6);`. Now you apply that function to the two lists $T[1]$ and $T[2]$, one element at a time, creating a list P of points as a result:

```
P := zip(pair,T[1],T[2]);.
```

The function *zip* takes a binary function and two lists as arguments and processes the two lists, applying the functions to the corresponding components of the two lists, one component at a time. See the Help information for *zip* for more information.

Now you can plot the data: `plot(P);`. Try right clicking on the plot and changing the **Line Style** or **Legend**.

Exercise 6

Comparing the different curves. Change the equations for the curves you found above to functions of the variable x. One way to do this is to use the functions *unapply* and *rhs*, for example with `lineq:=unapply(rhs(lin),x);`. This makes a function of x out of the right-hand side (rhs) of *lin*. Do this for the other two curves and call the functions xlog and xlog2. Now try plotting the original points and these three functions to see which looks best. You can use

```
plot([P,lineq(x),xlog(x),xlog2(x)],x=1..8,y=0..40,color=[red,green,blue,plum]);.
```

Try leaving out the restrictions on x and y to see why they are included. Often you have to try out different possibilities to see what gives you the most information. There seems to be a difference between the functions at the value 3, so try evaluating all three functions at 3. For example, try `lineq(3);` and `T[2,3];`. With some of the results, you should try approximating to five digits precision.

Exercise 7 – Optional –Do if you have time

You can create lists of the values of your functions at the same values of x as for your original points. Look at these original values `T[1];`. Then try mapping your functions *lineq*, *xlog*, and *xlog2* onto every element of this, for example `linlist:=map(lineq,T[1]);`. You can right click on the lists to get approximations of all of the values.

Another way to compare these curves is to create histograms, though the results will not have as much detail as with the other plots. Try

```
histogram(T[2],linlist,area=count,numbars=8,color=cyan);.
```

Try right clicking on the histogram and changing the color scheme. The result is 3-dimensional, so you can rotate it. Find the symbols ϑ and φ in the line below the toolbar, and click on the arrows to change the angles for viewing the histogram. You can also plot all four sets of points in one histogram, by giving four lists, instead of just two.

Exercise 8

Using a spreadsheet (regneark) in Maple.

The spreadsheets in Excel and Maple are compatible, in that you can copy a spreadsheet from Excel over to Maple. The formulas will disappear and you will just get the data. Many of the same ideas from the Excel spreadsheets also apply to Maple, but the Maple spreadsheets are much less convenient for the standard bookkeeping work that Excel is so good at. On the other hand, it is better at creating sequences of mathematical formulas. The following will introduce you to spreadsheets in Maple.

To get started you need to get the Maple 7 capabilities into Excel. Do this as explained on

http://www.imada.sdu.dk/~joan/IT/ExcelMaple7.html

Easier version of exercise 8

First, you will copy data from an Excel spreadsheet to a Maple spreadsheet. Find the Excel file where you stored the linear regression data you just finished working with. When you double click on it, Excel will start up with the file. Select the two columns of data and click on the first of the Maple buttons on the toolbar (it is red and white, looks like it should be used for copying, and says "Copy From Excel to Maple" when you move your mouse to it). Close or iconify the Excel window. In Maple, start with the command `with(Spread);`. The package `Spread` is the spreadsheet package. Then, choose **Spreadsheet** from the **Insert** menu button. This should produce a spreadsheet. You can make it bigger by dragging the lower right hand corner. In cell $A1$ type `n`, and in cell $B1$ type `f(~A1)`. Notice that in a Maple spreadsheet, you reference the contents of a cell by placing a `~` before the name of the cell. Select cell $A2$. You can use either the toolbar button or the **Edit** menu to *Paste* the two columns of data in.

If you right click on the spreadsheet, you can click on **Properties...** and find the spreadsheet's name and change it to something shorter, like *regress*. If you wish, you can make other changes, such as changing the colors of some cells. Then, click on **Apply** and **OK**.

In column C you can easily get the results of applying the linear approximation to the values 1 through 20. In cell $C1$, type `lineq(~A1)`. Now select cells $C1$ through $C21$. From the **Spreadsheet** menu, choose **Fill** and **Down**. You should have the result now.

One method for getting your information out of the spreadsheet into some of the Maple data types you are used to is the following: Create a new *Execution Group* below your spreadsheet. Delete the first row in your spreadsheet by selecting it and using the **Spreadsheet** menu. Choose **Execute** and **Worksheet** from the **Edit** menu. Select all the data remaining in your spreadsheet. Type `GetValuesMatrix(regress);` (assuming you named your spreadsheet *regress*). Right click on the output and choose **Transpose** (which will, of course, transpose the matrix). Assuming that the output of this is called $R0$, type `V:=convert(R0,listlist);` to change from a matrix to a list of lists. Now you should see all of the data from the spreadsheet. Type `V[3];` to see the third column of data from the spreadsheet. You can use these lists exactly as you did earlier.

Harder version of exercise 8

Start by copying the first $1/s$ and $1/v$ data from your first Excel lab. Remember to have headers for your data. Copying this data is harder than copying the integer data from your other regression test, since the Danish version of Excel uses commas and Maple does not recognize them properly. To avoid this problem, before copying, change your Excel numbers to fractions with three digits of precision. (Remember to change them back after you do the copy. You can use **Undo** from the **Edit** menu.) Then do the copy from Excel to Maple. Within the Maple spreadsheet, you can change the properties of those columns by selecting them and right clicking, choosing **Properties** and changing to floating point. You should also give the spreadsheet a reasonable name by changing its **Properties**.

Create a new execution paragraph before your spreadsheet and define a function to take the inverses of your data (to get back the original data from Excel lab). In the third and fourth columns compute the inverses of your first and second columns, respectively.

Now get the data out of your spreadsheet, first into a matrix and then into a list of lists.

Comments for both versions of exercise 8

More generally, you can always get data from Excel to Maple in this way (though, there is the restriction that Maple's spreadsheets have size no greater than 100 by 52, so you may have to use more than one spreadsheet if you have a lot of data). It is also possible to get data from Maple to Excel, similarly. Actually, you can skip using the spreadsheet in Maple when copying data from Excel to Maple. If you click on an execution paragraph, rather than in a spreadsheet, the data will be put in a matrix.

Exercise 9

Save your worksheet. E-mail your worksheet to your lab instructor as an attachment. Remember to logoff your computer, but do not push any of the buttons on it.