MS Excel and VBA

Module 1: Highlights of Excel

Bruno Abreu Calfa

Last Update: November 16, 2011

Table of Contents

Outline

Contents

1 Introduction and Requirements

Introduction

- Microsoft Excel is a *spreadsheet* software that is part of Microsoft Office suite
- Spreadsheet application means: multiple cells usually in a two-dimensional matrix or grid consisting of rows and columns
- Excel's main features are: calculation, graphing, and a "macro" programming language called *Visual Basic for Applications* (VBA)
- VBA allows you to programmatically operate on the information stored in Microsoft Office applications (not only in Excel)

Software and Programming Requirements

- The material covered here requires either Microsoft Excel 2010 for Windows or Excel for Mac 2011
- Even though both versions support VBA, Excel for Windows is used in the following slides
- Virtually all functions and options available in Excel for Windows are also found in Excel for Mac
- Some experience with Excel is assumed (this is not a comprehensive introduction to Excel)
- Previous programming experience will be helpful, but no experience with VBA is assumed

Help with Excel and VBA

- Google Observation: Many times I find answers at www.ozgrid.com
- References used in this material:
  - Walkenbach, J. (2010) **Excel® 2010 Power Programming with VBA**. Wiley Publishing, Inc. 1080p.
  - Billo, E. J. (2007) **Excel® for Scientists and Engineers: Numerical Methods**. John Wiley & Sons, Inc. 480p.
  - Webb, J., Saunders, S. (2006) **Programming Excel with VBA and .NET**. O'Reilly Media. 1120p.

2 Basics of Excel

Writing in a Worksheet

- Hierarchy of *objects*: `Application` → `Workbook` → `Worksheet` → `Range` → `Cell`
- Each cell can contain text, numbers, and formulas
- Formulas begin with an equal sign, =
- To "display" a formula without evaluating it, begin with a single quote, `'`
- Use the many formatting options from the tab "Home" in the ribbon to change fonts, colors, add borders *etc*.
- See file **`Excel_Basics.xlsx`**, worksheet "Writing in a Worksheet"

Using the Fill Handle

- *Fill handle* is the black square that appears at the bottom right corner of an *active* cell (when it is selected)
- When dragged, it fills other cells with the same content of the original cell or *intelligently* fills them with other contents
- For numeric cells, hold the `CTRL` key to increase the number by a unit (notice the cursor changes its shape)
- For alphanumeric cells terminated in a number, drag the fill handle to increment the numeric part by a unit
- For adjacent columns, double-click the fill handle to autocomplete the remaining cells of the new column (works for formulas too)
- Best explained with examples (see file **`Excel_Basics.xlsx`**, worksheet "Fill Handle")

3 Some Interesting Features/Functions

Conditional Formatting

- Excel allows you to format cells based on their values
- For example, apply the following conditional formatting:

Value	Color
$v \leq 4$	red
$4 \leq v \leq 7$	blue
$v \geq 7$	green

- The "Conditional Formatting" button is in the "Home" tab
- Choose the formatting *rules* and apply them to a given range or the entire worksheet
- See file **`Excel_Basics.xlsx`**, worksheet "Conditional Formatting"

Data Tables

- Insert data tables (with headers) to allow you to filter and sort data
- First enter the data in columns and then click on the "Table" button in the "Insert" tab
- Select the range that contains your data and, if it has headers, make sure the option is checked

- After the table is created, click on the downward arrow to the right of the header cells to sort and filter what is displayed
- See file **`Excel_Basics.xlsx`**, worksheet "Data" contains the raw data and worksheet "Data Table" contains the data table generated

Pivot Tables

- The pivot tables are used to summarize, analyze, explore and present your data
- First enter the data in columns and select all the data (including headers) that you want to include in the pivot table
- Then click on "Pivot Table" from the "Insert" tab
- Choose the destination of the pivot table (*e.g.* New Worksheet)
- See file **`Excel_Basics.xlsx`**, worksheet "Data" contains the raw data and worksheet "Pivot Table" contains the pivot table generated

Lookup Functions

- Two very useful Worksheet functions: `VLOOKUP` and `HLOOKUP`
- They retrieve information from a database/list based on a supplied instance of the unique identifier (your data do not need to be in a Table to use these functions)
- The `V` in `VLOOKUP` stands for "vertical", which means that the comparison values are located in *columns*
- The basic syntax for `VLOOKUP` is: given a unique value, the data range, the column number of the resulting value, and a flag to denote if the column containing the unique value is sorted, retrieve the corresponding value to that instance of the unique identifier
- See file **`Excel_Basics.xlsx`**, worksheet "Lookup"

4 Applications: Lookup and Roots of Equations

Two-Way Table Lookup

- Combine the Worksheet functions `VLOOKUP` and `MATCH` to obtain a value from a two-way table
- The function `MATCH` returns the index of a value in an array or #N/A if not found
- Idea: pass the row value, the table range, and the result of `MATCH` to the function `VLOOKUP`
- Probably will have to add 1 to the result of `MATCH` to account for the offset of one column
- See file **`Excel_Applications.xlsx`**, worksheet "2-Way Lookup"
- Notice that we are **not** performing *interpolation*

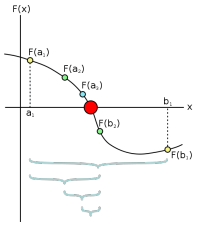
Finding Roots of Equations

- Given the general nonlinear equation
$$f(x) = 0$$
the goal is to find a value x^* such that $f(x^*) = 0$
- There are sophisticated ways to numerically find roots of equations (one can use the Solver tool for that)
- Let us look at three (iterative) methods

4.1 Derivative-Free Methods

Bisection Method

- It is based on the *Intermediate Value Theorem*
- Given an interval $[a,\ b]$ over which $f(x)$ is defined
- If $f(a)$ and $f(b)$ have opposite *signs*, then there is a root between a and b

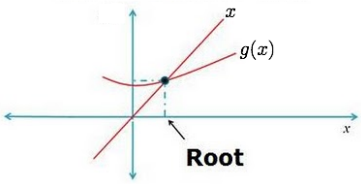


- Idea: compute the intermediate point $c_k = (a_k + b_k)/2$ and decide the next values of the bounds a_{k+1} and b_{k+1} based on the signs of $f(c_k)$, $f(a_k)$, and $f(b_k)$
- Stop after some tolerance has been reached
- See file **`Excel_Applications.xlsx`**, worksheet "Bisection"

Successive Substitution Method

- It is based on the *fixed-point* problem

$$f(x) = x - g(x) = 0$$



- Iteration formula:

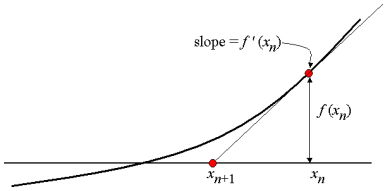
$$x_{k+1} = g(x_k)$$

- Idea: start from initial guess x_0, compute $g(x_0)$ and check if $f(x_0)$ is sufficiently close to 0; if not, use iteration formula until some tolerance is reached
- See file **`Excel_Applications.xlsx`**, worksheet "Substitution"

4.2 Method that uses Derivative Information

Newton's Method

- Also called Newton-Raphson Method



- Iteration formula:

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

- Idea: start from initial guess x_0, compute $f(x_0)$, $f'(x_0)$, and x_1 and check for convergence according to some tolerance criterion(a)
- See file **`Excel_Applications.xlsx`**, worksheet "Newton"