

Scientific and Mathematical Computing Using Python

Adam Cunningham

University at Buffalo
Department of Biostatistics

This work is licensed under the Creative Commons Attribution 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/> or send a letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.

Contents

| | | |
|-----------|---|-----------|
| I | Getting Started | 7 |
| 1 | Running Python | 8 |
| 1.1 | Install Python | 8 |
| 1.2 | Jupyter Notebook | 9 |
| 1.2.1 | The "Files" Tab | 9 |
| 1.2.2 | The "Running" Tab | 9 |
| 1.2.3 | The "Clusters" Tab | 9 |
| 1.2.4 | The "Conda" Tab | 9 |
| 1.2.5 | Starting a New Notebook | 10 |
| 1.2.6 | Magics | 12 |
| 1.2.7 | Markdown | 13 |
| 1.2.8 | L ^A T _E X | 14 |
| II | Programming Python | 15 |
| 2 | Python Data Types | 16 |
| 2.1 | Numbers | 16 |
| 2.2 | Booleans | 18 |
| 2.3 | Strings | 18 |
| 2.4 | Formatting Strings | 21 |
| 2.5 | Type Conversions | 22 |
| 2.6 | Variable Names | 24 |
| 3 | Containers | 26 |
| 3.1 | Lists | 26 |
| 3.2 | Tuples | 29 |
| 3.3 | Sets | 30 |
| 3.4 | Dictionaries | 32 |
| 4 | Controlling the Flow | 34 |
| 4.1 | Boolean Expressions | 34 |
| 4.2 | If Statements | 34 |
| 4.3 | Conditional Expressions | 36 |

| | | |
|------------|-------------------------------------|-----------|
| 4.4 | For Loops | 37 |
| 4.5 | While Loops | 40 |
| 4.6 | Break and Continue | 40 |
| 4.7 | Error Handling with Try-Except | 41 |
| 4.8 | Reading and Writing Files | 42 |
| 5 | Packaging and Reusing Code | 45 |
| 5.1 | Functions | 45 |
| 5.2 | Modules | 48 |
| 5.3 | Comprehensions | 49 |
| 5.4 | Generator Expressions | 50 |
| 5.5 | Comments | 50 |
| III | Numerical Computing | 52 |
| 6 | NumPy | 53 |
| 6.1 | Array Creation | 54 |
| 6.2 | Array Properties | 57 |
| 6.3 | Array Operations | 58 |
| 6.4 | Array Indexing and Slicing | 60 |
| 6.5 | Indexing with Integer Arrays | 61 |
| 6.6 | Indexing with Boolean Arrays | 63 |
| 7 | Matplotlib | 65 |
| 7.1 | Basic Plotting | 66 |
| 7.2 | A More Complex Plotting Example | 67 |
| 7.3 | Bar Plots | 68 |
| 7.4 | Polar Plots | 69 |
| 7.5 | Histograms | 70 |
| 7.6 | Pie Charts | 71 |
| 7.7 | Contour Plots | 72 |
| 7.8 | Slope Fields | 73 |
| 7.9 | Stream Plots | 74 |
| 7.10 | Multiple Plots | 75 |
| 7.11 | Formatting Text | 76 |
| 7.12 | Formatting Mathematical Expressions | 77 |
| 8 | Differential Equations | 78 |
| 8.1 | First-Order Differential Equations | 79 |
| 8.2 | Higher Order Linear Equations | 80 |
| 8.3 | Systems of Equations | 81 |
| 9 | Additional Topics | 82 |
| 9.1 | Loading Numerical Files | 82 |

| | | |
|-----------|--|-----------|
| 9.2 | Images | 83 |
| 9.3 | Animation | 84 |
| 9.4 | Random Number Generation | 86 |
| 9.5 | Sound Files | 87 |
| 9.6 | Linear Programming | 88 |
| IV | Programming Tips | 90 |
| 10 | Programming Style | 91 |
| 10.1 | Choosing Good Variable Names | 91 |
| 10.2 | Choosing Good Function Names | 92 |
| 10.3 | No “Magic Numbers” | 92 |
| 10.4 | Comments | 92 |
| 10.5 | Errors and Debugging | 93 |
| 11 | Further Reading | 94 |
| | Index | 95 |

Preface

This book covers material used in the courses "MTH 306: Differential Equations" and "MTH 337: Introduction to Scientific and Mathematical Computing" taught at the University at Buffalo. The following areas are covered:

- Programming using Python, the scientific computing package NumPy, and the plotting library Matplotlib.
- Python packages available for solving differential equations, random number generation, linear programming, animation, and loading numerical, image, and sound data.
- Programming tips relating to good programming style.

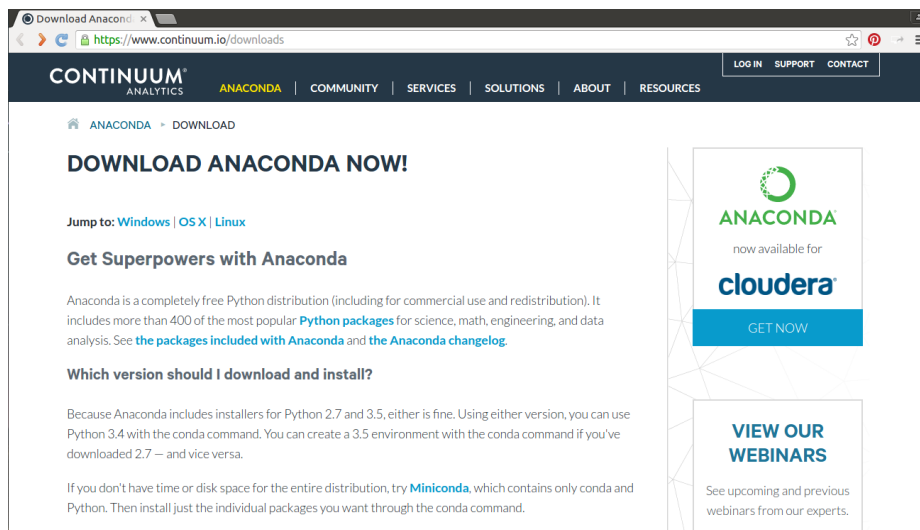
Python is a beautiful language. Have fun!

1 Getting Started

I Running Python

II Install Python

The examples in this book use Python 3.5. It is recommended that you use the [Anaconda](#) distribution, which is available free on Windows, Mac and Linux and contains all the packages needed (NumPy, SciPy, Matplotlib, Jupyter Notebook).



1.2 Jupyter Notebook

The development environment used in this book is the [Jupyter Notebook](#). This provides:

- An interactive environment for writing and running code.
- A way to integrate code, text and graphics in a single document.

1.2.1 The "Files" Tab

On starting Jupyter, a window with several tabs is displayed. The first of these is the "Files" tab. This provides a view of all the files in the folder where Jupyter Notebook was started, similar to a folder view in "Windows Explorer".



Files can be selected in this window then opened, duplicated, renamed or deleted. The [New](#) button can be used to create a new text file, folder, terminal, or Jupyter Notebook.

1.2.2 The "Running" Tab

The "Running" tab shows the currently active terminals or notebooks in Jupyter, and can be used to close them down.



1.2.3 The "Clusters" Tab

The third tab is the "Clusters" tab, which we will not consider in this class.



1.2.4 The "Conda" Tab

The "Conda" tab is used to manage the add-on packages that are available for Python. These provide additional functionality not present in the core system.

1 Conda environment

| Action | Name | Default? | Directory |
|--------|------|----------|----------------------|
| | root | ✓ | /home/adam/anaconda3 |

Available packages

Search...

198 installed packages in environment "root"

| Name | Version | Channel | Available |
|---|---------|---------|-----------|
| <input type="checkbox"/> _license | 1.1 | py35_1 | |
| <input type="checkbox"/> _nb_ext_conf | 0.3.0 | py35_0 | |
| <input type="checkbox"/> alabaster | 0.7.9 | py35_0 | |
| <input type="checkbox"/> anaconda | custom | py35_0 | |
| <input type="checkbox"/> anaconda-client | 1.6.0 | py35_0 | |
| <input type="checkbox"/> anaconda-navigator | 1.3.2 | py35_0 | |

Installing a new package is done by selecting it from the list of available packages in the left-hand window, then clicking the button to install it. The package will be displayed in the right-hand window once it has been downloaded and installed.

1.2.5 Starting a New Notebook

A new Jupyter Notebook can be created from the "Files" tab using the button. The notebook opens in a web browser, and contains:

- A **Title** bar, containing the name of the notebook.
- A **Menu** bar, containing all the functions available in the notebook.
- A **Tool** bar, for easy access to the most commonly-used functions.
- A list of cells, containing code or text, and the results of executing the code.

Jupyter Untitled Last Checkpoint: a minute ago (unsaved changes)

File Edit View Insert Cell Kernel Widgets Help



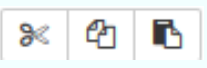




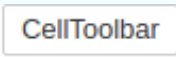


Python [default]

In []:


The menu bar contains the following functions.

| Menubar functions | |
|-------------------|--|
| Menu Item | Functions Available |
| File | Open, close, copy, save, or rename the current notebook. Export the notebook to a different format (HTML, PDF etc). |
| Edit | Cut, copy, paste or delete cells. Split and merge cells. Move cells up or down. Find and replace text in the notebook. |
| View | Turn the header, toolbar, or cell toolbars on or off. |
| Insert | Insert new cells into the notebook. |
| Cell | Run cells (i.e. execute the code they contain) Set the cell type (Code/Markdown) Turn cell output on or off. |
| Kernel | Interrupt or restart the Python "kernel". Interrupting the kernel is used if code is taking too long to execute. Restarting the kernel starts Python again in a clean state, without any additional variables or functions loaded. |
| Widgets | Embed widgets (controls for interactive data) |
| Help | Help with the Notebook, Python, and the most important add-on packages (NumPy, SciPy, Matplotlib). |

The tool bar contain the following functions.

| Toolbar functions | |
|---|--|
| Button | Function |
|  | Save current notebook |
|  | Insert a new cell below the currently selected one |
|  | Cut/copy/paste a cell |
|  | Move the selected cell up/down |
|  | Run cell/interrupt kernel/restart kernel |
|  | Set the cell type (usually Code or Markdown) |
|  | Open the palette of all cell commands |
|  | Open the cell toolbar |
|  | Publish this notebook to Anaconda.org |
|  | Edit/show a presentation based on this notebook |

Code and text are entered in *cells*, which can be of different types. The main types we will use are:

- **Code** cells, which contain Python code.
 - Click on a cell with the mouse to start entering code.
 - Enter adds a new line in the cell, without executing the code.
 - Shift-Enter (or clicking the "Play"  button in the toolbar, or **Cell** → **Run Cells** in the menubar) executes the code in the cell and moves the cursor to the next cell.
 - Tab brings up help for the function the cursor is currently in.
- **Markdown** cells contain text formatted using the Markdown language, and mathematical formulas defined using \LaTeX syntax.

The type can be selected by either using the "Cell Type" pull-down menu in the toolbar, or **Cell** → **Cell Type** in the menubar.

1.2.6 Magics

Magics are instructions that perform specialized tasks. They are entered and executed in code cells, and prefaced by "%" for a line magic (which just applies to one line) or "%%" for a cell magic (which applies to the whole cell). The main ones we will be using are:

- `%matplotlib inline` sets up the notebook so that plots are drawn inline (in the notebook itself).
- `%run <file>` executes the Python commands in `<file>`.
- `%timeit <code>` records the time it takes to run a line of Python code.
- `%%timeit` records the time it takes to run all the Python code in a cell.

An example of timing code execution using `%%timeit` is as follows,


| | |
|---|--|
| <pre>%%timeit x = range(10000) max(x)</pre> <hr/> <pre>1000 loops, best of 3: 884 μs per loop</pre> | <p><i>The line "x = range(10000)" is run once but not timed. The "max(x)" line is timed.</i></p> |
|---|--|



Note that the `%%timeit` magic *must* be the first line in the code cell.

1.2.1 Markdown

Text can be added to Jupyter Notebooks using [Markdown](#) cells. Markdown is a language that can be used to specify formatted text such as italic and bold text, lists, hyperlinks, tables and images. Some examples are shown below.

| Markdown | How it prints |
|---|---|
| # An h1 header | An h1 header |
| ## An h2 header | An h2 header |
| ### An h3 header | An h3 header |
| #### An h4 header | An h4 header |
| <i>*italic*</i> | <i>italic</i> |
| **bold** | bold |
| This is a bullet list * First item * Second item | This is a bullet list <ul style="list-style-type: none"> • First item • Second item |
| This is an enumerated list 1. First item 2. Second item | This is an enumerated list <ol style="list-style-type: none"> 1. First item 2. Second item |
| [UB link](https://www.buffalo.edu/) | UB link |
| ![Python](PythonImage.jpg "Python") |  |
| A horizontal line *** | A horizontal line <hr/> |

1.2.8 L^AT_EX

Mathematical expressions in Markdown cells are specified using the typesetting language L^AT_EX. These expressions are identified using `$(formula)$` for an inline formula (displayed within a line of text), or `$$ (formula) $$` for a larger formula displayed on a separate line.

| | |
|---|--|
| Superscripts | |
| <code>\$x^2\$</code> | x^2 |
| Subscripts | |
| <code>\$x_1\$</code> | x_1 |
| Fractions | |
| <code>\$\$\frac{1}{2}\$\$</code> | $\frac{1}{2}$ |
| Greek Letters | |
| <code>\$\$\alpha, \beta, \gamma, \dots, \omega\$</code> | $\alpha, \beta, \gamma, \dots, \omega$ |
| Series | |
| <code>\$\$\sum_{i=1}^n\$</code> | $\sum_{i=1}^n$ |
| Integrals | |
| <code>\$\$\int_a^b\$</code> | \int_a^b |
| Square Roots | |
| <code>\$\$\sqrt{a + b}\$</code> | $\sqrt{a + b}$ |
| Overline | |
| <code>\$\$\bar{x}\$</code> | \bar{x} |
| Brackets | |
| <code>\$\$\{1, 2, \dots, n\}\$</code> | $\{1, 2, \dots, n\}$ |
| Matrices | |
| <code>\$\$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\$</code> | $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ |

II Programming Python

2 Python Data Types

Python is a flexible and powerful high-level language that is well suited to scientific and mathematical computing. It has been designed with a clear and expressive syntax with a focus on ensuring that code is readable.

2.1 Numbers

The basic numerical types used in Python are:

- Integers.
- Floats (reals).
- Complex numbers (pairs of floats).

Python will automatically convert numbers from one type to another when appropriate. For example, adding two integers yields an integer, but adding an integer and a float yields a float. The main arithmetic operations are +, -, *, /, and **. Operations are evaluated in standard order - Parentheses, Exponentiation, Multiplication, Division, Addition, Subtraction. To avoid possible ambiguity, use parentheses to make the order of evaluation clear.

| | |
|-------------------------|-----------------------|
| $3 + 2$ ----- 5 | <i>Addition</i> |
| $3 - 2$ ----- 1 | <i>Subtraction</i> |
| $3 * 2$ ----- 6 | <i>Multiplication</i> |
| $3 / 2$ ----- 1.5 | <i>Division</i> |

| | |
|-------------------|--|
| <code>3**2</code> | <i>Exponentiation (not 3^2)</i> |
| 9 | |

Some other useful operations are floor division (`//`), modulus (`%`), and absolute value (`abs`).

| | |
|-----------------------|--|
| <code>3 // 2</code> | <i>Floor division returns the integral part of the quotient.</i> |
| 1 | |
| <code>12 % 5</code> | <i>Modulus returns the remainder.</i> |
| 2 | |
| <code>abs(-88)</code> | abs returns the absolute value. |
| 88 | |

Python has a built-in complex number type, and knows the rules of complex arithmetic.

| | |
|--|--|
| <code>1 + 2j</code> | <i>Generate a complex number (<i>j</i> is used instead of <i>i</i>).</i> |
| <code>(1+2j)</code> | |
| <code>complex(1, 2)</code> | <i>Another way to generate a complex number.</i> |
| <code>(1+2j)</code> | |
| <code>(1+2j).real</code> | real returns the real part of a complex number. |
| 1.0 | |
| <code>(1+2j).imag</code> | imag returns the imaginary part of a complex number. |
| 2.0 | |
| <code>abs(3+4j)</code> | abs returns the modulus when applied to a complex number. |
| 5.0 | |
| <code>z = 1 + 2j</code> <code>w = 3 - 1j</code> | <i>Note that a '1' is needed in front of the <i>j</i>.</i> |

| | |
|---------------------------------------|--------------------------------|
| <code>z + w</code> ----- (4+1j) | <i>Complex addition.</i> |
| <code>z * w</code> ----- (5+5j) | <i>Complex multiplication.</i> |

2.2 Booleans

Python also has a Boolean type, which only takes the values `True` or `False`. These can also be used like numbers, where `True` has the value 1 and `False` the value 0.

| | |
|--|---|
| <code>True or False</code> ----- <code>True</code> | <i>Logical disjunction</i> |
| <code>True and False</code> ----- <code>False</code> | <i>Logical conjunction</i> |
| <code>not True</code> ----- <code>False</code> | <i>Logical negation</i> |
| <code>True + 41</code> ----- 42 | <i>True has the numerical value 1.</i> |
| <code>False * 41</code> ----- 0 | <i>False has the numerical value 0.</i> |

2.3 Strings

Strings are sequences of characters. They are identified by surrounding quote marks.

- To generate a string, enclose a sequence of characters in either single (') or double (") quotes (Python doesn't care which).
- A single character in Python is just a one-element string.
- Python strings are *immutable* - once defined, they can't be changed. They can of course still be copied or operated on to create new strings.

| | |
|--|--|
| <pre>print("abc")</pre> abc | print outputs text to the screen (discarding the quotes). |
| <pre>"abc" + "def"</pre> "abcdef" | Adding two strings makes a new string by concatenation. |
| <pre>"abc"*3</pre> "abcabcabc" | Multiplying a string by an integer repeats the string. |
| <pre>print("I love 'MTH 337'!")</pre> I love 'MTH 337'! | Embedding quote marks within a string. |

A "\ " within a string is used to specify special characters such as newlines and tabs.

| | |
|--|---|
| <pre>string1 = "abc\ndef"</pre> <pre>print(string1)</pre> abc def | The "\n" character specifies a newline. |
| <pre>string2 = "abc\tdef"</pre> <pre>print(string2)</pre> abc def | The "\t" character specifies a tab. |

Strings elements are accessed using square brackets, [] .

- *Indexing* obtains characters from the string using a single integer to identify the position of the character.
- *Slicing* obtains a substring using `start:stop:step` to specify characters to select.
- Indexing and slicing are *zero-based* - the first character is at position 0.
- Indexing and slicing is "up to but not including" the stop position.
- A ":" can be used to select *all* characters either before or after a given position.

| Indexing the string "abcde" | | | | |
|-----------------------------|---|---|---|---|
| a | b | c | d | e |
| 0 | 1 | 2 | 3 | 4 |

| | |
|--|--|
| <code>"abcde"[1]</code> "b" | Indexing the character at index 1 (indices start at 0, not 1). |
| <code>"abcde"[-1]</code> "e" | Negative indices count backwards from the end of the string. |
| <code>"abcde"[1:4]</code> "bcd" | Slicing a string from position 1 up to (but not including) position 4. |
| <code>"abcde"[2:]</code> "cde" | Select all characters from position 2 to the end of the string. |
| <code>"abcde"[:2]</code> "ab" | Select all characters from the start of the string up to (but not including) position 2. |
| <code>"abcde"[: :2]</code> "ace" | Select every second character from the whole string. |
| <code>"abcdefg"[1:5:2]</code> "bd" | Select every second character from positions 1 up to 5. |
| <code>"abcde"[: :-1]</code> "edcba" | Reversing a string by reading it backwards. |

String *methods* are commands specific to strings that carry out some action on the string. They are called by following the string with a '.' followed by the method name.

| | |
|--|---|
| <code>"mth337".upper()</code> "MTH337" | upper converts all letters to uppercase. |
| <code>"ABCDE".lower()</code> "abcde" | lower converts all letters to lowercase. |
| <code>"mth337".capitalize()</code> "Mth337" | capitalize capitalizes the first letter. |

| | |
|---|--|
| <pre>"abcabcabc".replace("c", "z")</pre> <hr/> <pre>"abzabzabz"</pre> | replace replaces all occurrences of one substring with another substring. |
| <pre>"I love MTH337!".split()</pre> <hr/> <pre>["I", "love", "MTH337!"]</pre> | split splits a string into a list of words. |

String methods can also be used for counting the number of times a given substring occurs and for finding where it occurs.

| | |
|--|--|
| <pre>"avada kedavra".count("a")</pre> <hr/> <pre>5</pre> | count counts the number of times an item occurs in a list. |
| <pre>"avada kedavra".index("ked")</pre> <hr/> <pre>6</pre> | index finds the index where a substring occurs in a string. |

2.4 Formatting Strings

Strings can be formatted using the **format** function. This allows "replacement fields" surrounded by curly brackets `{}` in a string to be replaced by some other data. "Format specifications" within replacement fields define how data is to be formatted, including the field width, padding and number of decimal places.

- Empty replacement fields `{}` are filled in order with the arguments given.
- Numbers inside replacement fields specify arguments by position, starting with zero.

| | |
|---|---|
| <pre>"{} {}".format("a", "b")</pre> <hr/> <pre>"a b"</pre> | "Replacement fields" <code>{}</code> are filled in order by format . |
| <pre>"1st: {0}, 2nd: {1}".format(3,4)</pre> <hr/> <pre>"1st: 3, 2nd: 4"</pre> | The arguments to format can also be identified by position, starting at 0. |

Format specifications for the field width and padding are provided after a colon `:`.

- A field width can be specified using `:n` where `n` is an integer specifying the number of characters in the field.
- If the data has less characters than the field width, the default is to insert extra spaces on the right i.e. to "pad" on the right.

- To pad on the left, use `:>n`. To explicitly pad on the right, use `:<n`. The `>` and `<` are like direction arrows "sending" the data to the indicated side of the field.
- Data can be centered in a replacement field using `:^n`.

| | |
|--|--|
| <pre>print("{:5} {}".format("a", 2))</pre> <hr/> <pre>a 2</pre> | <p><i>Using a field of width 5 with the default padding.</i></p> |
| <pre>print("{:>5} {}".format("a", 2))</pre> <hr/> <pre> a 2</pre> | <p><i>Note the left padding in this example.</i></p> |
| <pre>print("{:^5} {}".format("a", 2))</pre> <hr/> <pre> a 2</pre> | <p><i>Now centering 'a' in the field of width 5.</i></p> |

Format specifications for floats allow tables of numerical data to be neatly printed and aligned.

- Integers are referred to using `:d`.
- Floats are referred to using `:f`.
- An integer after a decimal point is used to indicate how many decimal places to display.
- Use `:n.mf` to indicate a replacement field of width `n` for a float printed to `m` decimal places.

| | |
|--|---|
| <pre>print("{:5.2f}".format(pi))</pre> <hr/> <pre>3.14</pre> | <p><i>Print pi in a field of width 5 to 2 decimal places</i></p> |
| <pre>print("{:^10.4f}".format(pi))</pre> <hr/> <pre>3.1416</pre> | <p><i>Padding can be combined with other options if the padding is specified first.</i></p> |

2.5 Type Conversions

Objects can be explicitly converted from one type to another, as long as the conversion makes sense. This is called *type casting*.

- Ints can be cast to floats, and both ints and floats can be cast to complex numbers.
- Complex numbers can't be converted to ints or floats.

- Any object can be cast to a string.
- Strings can be cast to numerical types if the string represents a valid number.

Casting is done using the functions **bool**, **int**, **float**, **complex**, and **str**.

| | |
|--|--|
| <code>bool(1)</code> True | <i>Convert integer to boolean.</i> |
| <code>bool(42)</code> True | <i>Any nonzero value counts as True.</i> |
| <code>bool(0)</code> False | <i>Zero equates to False.</i> |
| <code>bool("")</code> False | <i>An empty string is also False.</i> |
| <code>int(2.99)</code> 2 | <i>Convert float to integer (the decimal part is discarded).</i> |
| <code>int("22")</code> 22 | <i>Convert string to int.</i> |
| <code>float("4.567")</code> 4.567 | <i>Convert string to float.</i> |
| <code>complex("1+2j")</code> (1+2j) | <i>Convert string to complex.</i> |
| <code>float(10)</code> 10.0 | <i>Convert integer to float.</i> |
| <code>complex(10)</code> (10+0j) | <i>Convert integer to complex number.</i> |
| <code>str(True)</code> "True" | <i>Convert boolean to string.</i> |

| | |
|--|---|
| <pre>str(1)</pre> <hr/> <pre>"1"</pre> | <i>Convert integer 1 to string "1".</i> |
| <pre>str(1.234)</pre> <hr/> <pre>"1.234"</pre> | <i>Convert float to string.</i> |

2.6 Variable Names

Variable names can be used to refer to objects in Python. They:

- Must start with either a letter or an underscore.
- Are case sensitive. So `value`, `VALUE`, and `Value` all refer to different variables.
- Are assigned a value using `"="`. The variable name goes to the left of the `"="`, and the value to assign on the right.

| | |
|---|---|
| <pre>x = 5</pre> <pre>print(x)</pre> <hr/> <pre>5</pre> | <i>Assign <code>x</code> the value 5 (note that <code>"="</code> is used for assignment, not <code>"=="</code>).</i> |
| <pre>y = x + 3</pre> <pre>print(y)</pre> <hr/> <pre>8</pre> | <i>Assign <code>y</code> the value of <code>x + 3</code>.</i> |
| <pre>course = "MTH 337"</pre> <pre>print(course)</pre> <hr/> <pre>MTH 337</pre> | <i><code>course</code> is a string (printed without quotes).</i> |
| <pre>a, b = 2, 3</pre> <pre>print(a, b)</pre> <hr/> <pre>2 3</pre> | <i>Multiple variables can be assigned at the same time.</i> |
| <pre>a, b = b, a</pre> <pre>print(a, b)</pre> <hr/> <pre>3 2</pre> | <i>The values of <code>a</code> and <code>b</code> are swapped (note that the right hand side is evaluated completely before the assignment is done).</i> |

| | |
|---|--|
| <pre>z = 3 z += 2 print(z)</pre> <hr/> <p>5</p> | <p><i>Same as $z = z + 2$.</i></p> |
| <pre>z -= 1 print(z)</pre> <hr/> <p>4</p> | <p><i>Same as $z = z - 1$.</i></p> |
| <pre>z *= 3 print(z)</pre> <hr/> <p>12</p> | <p><i>Same as $z = z * 3$.</i></p> |
| <pre>z /= 2 print(z)</pre> <hr/> <p>6</p> | <p><i>Same as $z = z / 2$.</i></p> |
| <pre>z %= 5 print(z)</pre> <hr/> <p>1</p> | <p><i>Same as $z = z \% 5$.</i></p> |

3 Containers

We often need a way to group objects of a similar kind together. Sometimes the order of the objects is important, as in a mathematical sequence $\{a_0, a_1, a_2, \dots\}$, and sometimes the order is irrelevant, as in the set $\{\text{'cat'}, \text{'dog'}, \text{'hippogriff'}\}$.

Python provides several types of *container* that can be used to group objects together. Containers differ in the following ways:

- *Mutable versus immutable*. Mutable containers can be modified after they have been created; immutable containers cannot.
- *Ordered versus unordered*. The items in an ordered container are stored in a fixed sequence; those in an unordered container are not.
- *Indexable versus non-indexable*. The items in an indexed container can be retrieved using a key; those in a non-indexable container cannot.

3.1 Lists

A list is an ordered sequence of objects, identified by surrounding square brackets, `[]`.

- To generate a list, enclose a sequence of objects (separated by commas) in square brackets.
- List elements can be of any type, and can be of different types within the same list.
- Lists are *mutable* - once created, elements can be added, replaced or deleted.

```
mylist = [1, "a", 6.58]  
print(mylist)
```

```
[1, "a", 6.58]
```

Use square brackets to create a list.

```
len(mylist)
```

```
3
```

len returns the number of elements in a list.

| | |
|--|---|
| <pre>list1 = [1, 2, 3] list2 = [4, 5, 6] list1 + list2</pre> <hr/> <pre>[1, 2, 3, 4, 5, 6]</pre> | <p><i>Adding two lists makes a new list by concatenation.</i></p> |
| <pre>list1 * 3</pre> <hr/> <pre>[1, 2, 3, 1, 2, 3, 1, 2, 3]</pre> | <p><i>Multiplying a list by an integer repeats the list.</i></p> |
| <pre>list3 = [] print(list3)</pre> <hr/> <pre>[]</pre> | <p><i>list3 is an empty list.</i></p> |
| <pre>list4 = list() print(list4)</pre> <hr/> <pre>[]</pre> | <p><i>Another way to create an empty list.</i></p> |

Lists can be indexed and sliced in the same way as strings, using square brackets. Indexing and slicing can also be used with the "=" assignment operator to change the elements of a list.

| | |
|---|--|
| <pre>primes = [2, 3, 5, 7, 11, 13, 17] primes[1]</pre> <hr/> <pre>3</pre> | <p><i>Access the element at index 1 (indexing starts at 0).</i></p> |
| <pre>primes[3:]</pre> <hr/> <pre>[7, 11, 13, 17]</pre> | <p><i>List slicing, start at position 3, through to the end.</i></p> |
| <pre>primes[:3]</pre> <hr/> <pre>[2, 3, 5]</pre> | <p><i>List slicing, start at the beginning, end at position 2.</i></p> |
| <pre>primes[2:5]</pre> <hr/> <pre>[5, 7, 11]</pre> | <p><i>List slicing, start at position 2, end at position 4.</i></p> |
| <pre>primes[::-1]</pre> <hr/> <pre>[17, 13, 11, 7, 5, 3, 2]</pre> | <p><i>One way to reverse a list.</i></p> |

| | |
|---|---|
| <pre>mylist[1] = "e" print(mylist)</pre> <hr/> <pre>["a", "e", "c"]</pre> | <p><i>List elements can be changed by assigning a new element at a given index.</i></p> |
|---|---|

List methods exist for changing, adding and deleting elements. These all modify an existing list.

| | |
|---|--|
| <pre>mylist = ["a", "b"] mylist.append("c") print(mylist)</pre> <hr/> <pre>["a", "b", "c"]</pre> | <p>append adds an element to the end of a list.</p> |
| <pre>mylist.extend(["d", "e", "f"]) print(mylist)</pre> <hr/> <pre>["a", "b", "c", "d", "e", "f"]</pre> | <p>extend adds elements to a list.</p> |
| <pre>mylist.insert(0, "z") print(mylist)</pre> <hr/> <pre>["z", "a", "b", "c", "d", "e", "f"]</pre> | <p>insert inserts an element at a given position.</p> |
| <pre>mylist.pop() print(mylist)</pre> <hr/> <pre>["z", "a", "b", "c", "d", "e"]</pre> | <p>pop removes the element at a given position and returns it. The default is to remove the last element.</p> |
| <pre>mylist.remove("z") print(mylist)</pre> <hr/> <pre>["a", "b", "c", "d", "e"]</pre> | <p>remove removes the first instance of an item from a list.</p> |

Methods also exist for counting items in a list and where they occur.

| | |
|---|--|
| <pre>fib = [1, 1, 2, 3, 5, 8, 13] fib.count(1)</pre> <hr/> <pre>2</pre> | <p>count counts the number of times an item occurs in a list.</p> |
| <pre>fib.index(13)</pre> <hr/> <pre>6</pre> | <p>index returns the index of the first occurrence.</p> |

Lists can be sorted and reversed.

| | |
|--|--|
| <pre>letters = ["a", "b", "c"] letters.reverse() print(letters)</pre> <hr/> <pre>["c", "b", "a"]</pre> | <p>reverse changes an existing list, reversing the order of elements.</p> |
| <pre>numbers = [2, 10, 3, 26, 5] print(sorted(numbers))</pre> <hr/> <pre>[2, 3, 5, 10, 26]</pre> | <p>sorted returns a sorted list, but does not modify the existing list.</p> |
| <pre>numbers.sort() print(numbers)</pre> <hr/> <pre>[2, 3, 5, 10, 26]</pre> | <p>sort sorts a list in place, modifying the existing list.</p> |
| <pre>sorted(numbers, reverse=True)</pre> <hr/> <pre>[26, 10, 5, 3, 2]</pre> | <p>The reverse keyword is used to sort in descending order.</p> |

The **min** and **max** functions find the smallest and largest items in a list.

| | |
|--|---|
| <pre>numbers = [2, 10, 3, 26, 5] print(min(numbers), max(numbers))</pre> <hr/> <pre>2 26</pre> | <p>min and max find the smallest and largest items.</p> |
|--|---|

3.2 Tuples

Tuples are containers like lists, with the difference being that they are *immutable* - once defined, elements cannot be changed or added. Tuples are identified by surrounding standard parentheses, `()`.

- To generate a tuple, enclose a sequence of objects (separated by commas) in standard parentheses.
- Tuple indexing and slicing works in the same way as for lists and strings.
- It is an error to try to change a tuple element once the tuple has been created.

Tuples are simpler and more efficient than lists in terms of memory use and performance. They are often preferred for "temporary" variables that will not need to be modified. They can also be used as dictionary keys, which lists cannot.

| | |
|--|---|
| <pre>tuple1 = ("a", "b", "c") print(tuple1)</pre> <hr/> <pre>("a", "b", "c")</pre> | <i>Create a tuple using standard parentheses.</i> |
| <pre>tuple1[2]</pre> <hr/> <pre>"c"</pre> | <i>Tuple elements can be indexed just like lists or strings.</i> |
| <pre>tuple1[1:]</pre> <hr/> <pre>("b", "c")</pre> | <i>Slicing works the same way for tuples as for lists or strings.</i> |

Any comma-separated sequence of values defines a tuple, which can be used to assign values to multiple variables at a time.

| | |
|---|---|
| <pre>tuple2 = 1, 2, 3 print(tuple2)</pre> <hr/> <pre>(1, 2, 3)</pre> | <i>A comma-separated sequence of values defines a tuple.</i> |
| <pre>(x, y) = (10, 20) print("x =", x) print("y =", y)</pre> <hr/> <pre>x = 10 y = 20</pre> | <i>The variables on the left-hand side are assigned to the values on the right.</i> |
| <pre>a, b = (2, 4) print(a, b)</pre> <hr/> <pre>2 4</pre> | <i>The parentheses are not strictly necessary, and can be discarded.</i> |

3.3 Sets

Sets are containers with the same meaning they do in mathematics - unordered collections of items with no duplicates. Sets are identified by surrounding curly brackets, {}.

- To generate a set, enclose a sequence of objects (separated by commas) in curly brackets.
- Duplicates will be removed when creating a set or operating on existing sets.
- Sets can be used instead of lists when we know that each element is unique and immutable (unchanging).

| | |
|---|---|
| <pre>myset = {1, 2, 3} print(myset)</pre> <hr/> <pre>set([1, 2, 3])</pre> | <i>Sets are created using curly brackets.</i> |
| <pre>myset = set([1, 2, 3, 2]) print(myset)</pre> <hr/> <pre>set([1, 2, 3])</pre> | <i>Creating a set from a list (note that duplicates are removed).</i> |
| <pre>print(set())</pre> <hr/> <pre>set([])</pre> | <i>set([]) creates an empty set.</i> |

The standard mathematical operations for sets are all built into Python.

| | |
|---|---|
| <pre>set1 = {1, 2, 3} set2 = {3, 4, 5}</pre> <hr/> | <i>Create 2 sets.</i> |
| <pre>1 in set1</pre> <hr/> <pre>True</pre> | in tests for set membership. |
| <pre>set1 set2</pre> <hr/> <pre>{1, 2, 3, 4, 5}</pre> | <i>Set union (the union operator can also be used).</i> |
| <pre>set1 & set2</pre> <hr/> <pre>{3}</pre> | <i>Set intersection (can also use the intersection operator).</i> |
| <pre>set1 - set2</pre> <hr/> <pre>{1, 2}</pre> | <i>Set difference (can also use the difference operator).</i> |
| <pre>set1 ^ set2</pre> <hr/> <pre>{1, 2, 4, 5}</pre> | <i>Symmetric difference (can also use the symmetric_difference operator).</i> |
| <pre>set1 <= set2</pre> <hr/> <pre>False</pre> | <i>Test if one set is a subset of another (can also use the is-subset operator).</i> |

3.4 Dictionaries

Dictionaries are containers where items are accessed by a key. This makes them different from *sequence* type objects such as strings, lists, and tuples, where items are accessed by position.

- To generate a dictionary, enclose a sequence of `key:value` pairs (separated by commas) in curly brackets.
- The key can be any immutable object - a number, string, or tuple.
- New dictionary elements can be added, and existing ones can be changed, by using an assignment statement.
- Order is *not* preserved in a dictionary, so printing a dictionary will not necessarily print items in the same order that they were added.

| | |
|--|--|
| <pre>dict1 = {"x":1, "y":2, "z":3} print(dict1) {"x": 1, "y": 2, "z": 3}</pre> | <p><i>Note the colon in the key:value pairs.</i></p> |
| <pre>dict1["y"] 2</pre> | <p><i>Dictionary values are accessed using the keys</i></p> |
| <pre>dict1["y"] = 10 print(mydict) {"x": 1, "y": 10, "z": 3}</pre> | <p><i>Dictionary values can be changed using the "=" assignment operator.</i></p> |
| <pre>dict1["w"] = 0 print(mydict) {"x": 1, "y": 10, "z": 3, "w": 0}</pre> | <p><i>New key:value pairs can be assigned using the "=" assignment operator.</i></p> |
| <pre>dict1.get("a")</pre> | <p>get returns <i>None</i> if the key does not exist.</p> |
| <pre>dict1.get("a", 42) 42</pre> | <p>get can also return a default value.</p> |
| <pre>dict2 = {} print(dict2) {}</pre> | <p><i>Creating an empty dictionary.</i></p> |


```
dict3 = dict()
print(dict3)
{}
```

Another way to create an empty dictionary.



It is an error to attempt to access a dictionary using a key that does not exist. This can be avoided by using the **get** method, which returns a default value if the key is not found.

4 Controlling the Flow

4.1 Boolean Expressions

Boolean expressions are statements that either evaluate to `True` or `False`. An important use of these expressions is for tests in *conditional code* that only executes if some condition is met. Examples of Boolean expressions include the standard comparison operators below.

| | |
|-----------------------------------|--|
| <code>5 == 5</code> True | <i>Check for equality.</i> |
| <code>5 != 5</code> False | <i>Check for inequality.</i> |
| <code>3 < 2</code> False | <i>Less than.</i> |
| <code>3 <= 3</code> True | <i>Less than or equals.</i> |
| <code>"a" < "b"</code> True | <i>Strings are compared by lexicographic (dictionary) order.</i> |

Note that *any* empty container evaluates to `False` in a Boolean expression. Examples include empty strings (`""`), lists (`[]`), and dictionaries (`{}`).

4.2 If Statements

Python **if** statements provide a way to execute a block of code only if some condition is `True`.

| if statement |
|---|
| <pre>if <condition>: <code to execute when condition True> <following code></pre> |

Note that:

- `<condition>` is a Boolean expression, which must evaluate to `True` or `False`.
- `<condition>` must be followed by a colon, `:`.
- The block of code to execute if `<condition>` is `True` starts on the next line, and *must be indented*.
- The convention in Python is that code blocks are indented with *4 spaces*.
- The block of code to execute is finished by de-indenting back to the previous level.

```
from math import *
if pi > e:
    print("Pi is bigger than e!")
```

```
Pi is bigger than e!
```

*The block of code following the **if** statement only executes if the condition is met.*

An **else** statement can be added after an **if** statement is complete. This will be followed by the code to execute if the condition is `False`.

if-else statement

```
if <condition>:
    <code to execute when condition True>
else:
    <code to execute when condition False>
<following code>
```

The following example illustrates an **if-else** statement.

```
x, y = 2**3, 3**2
if x < y:
    print("x < y")
else:
    print("x >= y")
```

```
x < y
```

*The block of code following the **else** statement executes if the condition is not met.*

Multiple **elif** statements (short for else-if) can be added to create a series of conditions that are tested in turn until one succeeds. Each **elif** must also be followed by a condition and a colon.

elif statement

```
if <condition 1>:
    <code to execute when condition 1 True>
elif <condition 2>:
    <code to execute when condition 2 True>
else:
    <code to execute if neither condition is True>
<following code>
```

The following example illustrates a series of conditions being tested.

```
score = 88
if score >= 90:
    print("A")
elif score >= 80:
    print("B")
elif score >= 70:
    print("C")
elif score >= 60:
    print("D")
else:
    print("F")
```

B

Only the first two conditions are tested - the rest are skipped since the second condition is True.

4.3 Conditional Expressions

It often happens that we want to assign a variable name some value if a condition is True, and another value if a condition is False. Using an **if** statement, we would have:

```
if <condition>:
    x = <>true_value>
else:
    x = <>false_value>
```

Python provides an elegant way to do the same thing in a single line using a *conditional expression*.

Conditional expression

```
x = <>true_value> if <condition> else <>false_value>
```

An example is given below.

```
x = 22
parity = "odd" if x % 2 else "even"
print(x, "has", parity, "parity")
```

```
22 has even parity
```

Note that $x \% 2$ returns the remainder when x is divided by 2. Any nonzero value evaluates as `True`.

4.4 For Loops

Python **for** loops provide a way to iterate (loop) over the items in a list, string, tuple, or any other *iterable* object, executing a block of code on each pass through the loop.

for loop

```
for <iteration variable(s)> in <iterable>:
    <code to execute each time>
<following code>
```

Note that:

- The **for** statement must be followed by a colon, `:`.
- One or more *iteration variables* are bound to the values in `<iterable>` on successive passes through the loop.
- The block of code to execute each time through the loop starts on the next line, and must be indented.
- This block of code is finished by de-indenting back to the previous level.

Sequence objects, such as strings, lists and tuples, can be iterated over as follows.

```
for i in [2, 4, 6]:
    print(i)
```

```
2
4
6
```

Iterate over the elements of a list. The iteration variable `i` gets bound to each element in turn.

```
for char in "abc":
    print(char)
```

```
a
b
c
```

Iterate over the characters in a string. The iteration variable `char` gets bound to each character in turn.

| | |
|---|--|
| <pre>for i, char in enumerate("abc"): print(i, char)</pre> <hr/> <pre>0 a 1 b 2 c</pre> | <p>enumerate allows an iteration variable to be bound to the index of each item, as well as to the item itself.</p> |
|---|--|

The **range** function generates integers in a given range. It is often used inside a **for** loop to iterate over some sequence of integers.

- The start, stop and step parameters to **range** are similar to those used to slice lists and strings.
- Integers are only generated by **range** as needed, rather than as a list.

| | |
|---|--|
| <pre>for i in range(3): print(i)</pre> <hr/> <pre>0 1 2</pre> | <p>range(n) generates n consecutive integers, starting at 0 and ending at n - 1.</p> |
| <pre>teens = range(13, 20) print(list(teens))</pre> <hr/> <pre>[13, 14, 15, 16, 17, 18, 19]</pre> | <p>range(start, stop) generates consecutive integers, from start to stop - 1.</p> |
| <pre>evens = range(0, 9, 2) print(list(evens))</pre> <hr/> <pre>[0, 2, 4, 6, 8]</pre> | <p>The third step argument to range specifies the increment from one integer to the next.</p> |
| <pre>for i in range(5, 0, -1): print(i)</pre> <hr/> <pre>5 4 3 2 1</pre> | <p>range can also count backwards using a negative step size.</p> |
| <pre>total = 0 for i in range(1, 6): total += i print(total)</pre> <hr/> <pre>15</pre> | <p>Sum the numbers from 1 to 5.</p> |

| | |
|---|--|
| <pre>sum(range(1, 6))</pre> <hr/> <pre>15</pre> | <p><i>Another (simpler) way to sum the numbers in a given range.</i></p> |
|---|--|

Dictionary elements consist of key:value pairs. When iterated over, variables can be bound to the key, the value, or both.

| | |
|--|--|
| <pre>mydict = {"x":1, "y":2, "z":3} for key in mydict: print(key)</pre> <hr/> <pre>y x z</pre> | <p><i>Iteration over a dictionary binds to the key (note that order is not preserved in a dictionary).</i></p> |
|--|--|

| | |
|--|---|
| <pre>for value in mydict.values(): print(value)</pre> <hr/> <pre>2 1 3</pre> | <p><i>Use values to iterate over the dictionary values rather than the keys.</i></p> |
|--|---|

| | |
|---|---|
| <pre>for key, value in mydict.items(): print(key, value)</pre> <hr/> <pre>y 2 x 1 z 3</pre> | <p><i>Use items to iterate over the dictionary keys and values together.</i></p> |
|---|---|

We can also iterate in parallel over multiple lists of equal length by using **zip**. This generates a sequence of tuples, with one element of each tuple drawn from each list.

| | |
|--|---|
| <pre>courses = [141, 142, 337] ranks = ["good", "better", "best!"] zipped = zip(courses, ranks) print(list(zipped))</pre> <hr/> <pre>[(141, "good"), (142, "better"), (337, "best")]</pre> | <p><i>zip(courses, ranks) generates a sequence of tuples. Each tuple contains one course and one rank, with the tuples in the same order as the list elements.</i></p> |
|--|---|

| | |
|---|--|
| <pre>for course, rank in zipped: print(course, rank)</pre> <hr/> <pre>141 good 142 better 337 best!</pre> | <p><i>Multiple iteration variables can be bound at each iteration of a loop.</i></p> |
|---|--|

4.5 While Loops

Python **while** loops execute a block of code repeatedly as long as some condition is met.

while loop

```
while <condition>:  
    <code to execute repeatedly>  
<following code>
```

Note that for the loop to terminate, the code must change some part of the `<condition>` so that it eventually returns `False`.

```
i = 3  
while i > 0:  
    print(i)  
    i -= 1
```

```
3  
2  
1
```

The variable `i` is printed while it remains greater than zero. The code inside the loop must change the value of `i` to ensure that the loop eventually terminates.

4.6 Break and Continue

Sometimes we need to end a loop early, either by ending just the current iteration, or by quitting the whole loop. The statements **break** and **continue** provide a way to do this.

- To end the loop completely and jump to the following code, use the **break** statement.
- To end the current iteration and skip to the next item in the loop, use the **continue** statement. This can often help to avoid nested if-else statements.

```
vowels = "aeiou"  
for char in "jabbawocky":  
    if char in vowels:  
        print("First vowel is", char)  
        break
```

```
First vowel is a
```

*The **for** loop is terminated by **break** once the first vowel is found.*


```
total = 0
for char in "jabbawocky":
    if char in vowels:
        continue
    total += 1
print(total, "consonants found")
```

```
7 consonants found
```

*Skip over the vowels using **continue**, and just count the consonants.*

4.7 Error Handling with Try-Except

If an invalid operation is attempted when running code then an error is usually generated. Examples include dividing by zero, indexing past the end of a sequence, creating a floating point number that is too large, or adding two arrays of different sizes. In these circumstances, the code typically breaks at the point of the invalid operation.

A **try-except** statement can be used to handle errors more deliberately than having the code break. This allows us to first try to execute some code that we're not sure will work, and then execute some other code if it doesn't.

try-except statement

```
try:
    <code to attempt to execute>
except:
    <code to execute only when an error occurs>
```

Errors typically have a type associated with them, which specifies the kind of error that has occurred. Examples include `ZeroDivisionError`, `IndexError`, `OverflowError`, and `ValueError`. The "except" part of a try-except statement can be specialized to handle these different kinds of error.

try-except statement with named error types

```
try:
    <code to attempt to execute>
except ErrorType:
    <code to only execute when the named error occurs>
```

The following example shows how to gracefully handle an error resulting from division by zero.

```
for i in range(-2,3):
    print(10/i)
```

```
-5.0
-10.0
Traceback (most recent call last):
File "<stdin>", line 2, in <module>
ZeroDivisionError: division by zero
```

The number 10 is being divided by a sequence of integers, one of which happens to be zero. Without error handling, the code breaks when the zero is encountered and a "ZeroDivisionError" is raised.

```
for i in range(-2,3):
    try:
        print(10/i)
    except ZeroDivisionError:
        print("Can't divide by zero!")
```

```
-5.0
-10.0
Can't divide by zero!
10.0
5.0
```

*Using a **try-except** statement to handle the ZeroDivisionError allows the loop to run to completion without breaking.*

4.8 Reading and Writing Files

Several reports for this class will involve reading and analyzing data that has been stored in a file. This typically involves three steps:

- Open the file using the **open** function. This returns a *file handle* - an object we then use to access the text that the file contains.
- Process the file, either line-by-line, or as a single text string.
- Close the file. This is done using the **close** function.

It is possible to read in the entire contents of a file in one go using the functions **read** and **readlines**. However, we may not need to read the entire contents into memory if we are dealing with a large file and just want to extract some information from the text. In this case, it is preferable to iterate over the lines of text that the file contains.

The following examples assume that a file named "firstnames.txt" has been created in the directory that Python was started in. This file contains the three lines:

```
"firstnames.txt"
```

```
Leonard
Penny
Sheldon
```

| | |
|---|--|
| <pre>input_file = open("firstnames.txt")</pre> | <p>Open "firstnames.txt" for reading using open.</p> |
| <pre>for line in input_file: print(line, end='')</pre> <p>Leonard Penny Sheldon</p> | <p>The lines of an open file can be iterated over in a for loop. Note the use of the end keyword in <code>print(line, end='')</code>, since each line already ends with a new line.</p> |
| <pre>input_file.close()</pre> | <p>Close "firstnames.txt" using close.</p> |
| <pre>input_file = open("firstnames.txt") first_names = input_file.read() input_file.close() first_names</pre> <p>"Leonard\nPenny\nSheldon\n"</p> | <p>read reads in the whole file as a single string. The newlines at the end of each line are shown as "\n" characters.</p> |
| <pre>print(first_names)</pre> <p>Leonard Penny Sheldon</p> | <p>Printing a string causes the newline characters "\n" to be rendered as new lines.</p> |
| <pre>input_file = open("firstnames.txt") data = input_file.readlines() input_file.close() data</pre> <p>["Leonard\n", "Penny\n", "Sheldon\n"]</p> | <p>readlines reads in the whole file as a list, with each line as a separate string.</p> |

Files can also be opened for writing using the "w" option to **open**.

| | |
|---|--|
| <pre>data = ["Hofstadter", "?", "Cooper"] output_file = open("names.txt", "w") for name in data: output_file.write(name + "\n") output_file.close()</pre> | <p>Write each string in the <code>data</code> list to a separate line in the file. Note that new lines are not automatically included, so they need to be added.</p> |
|---|--|

```
input_file = open("names.txt")
last_names = input_file.read()
input_file.close()
print(last_names)
```

Hofstadter
?
Cooper

Check that the "names.txt" file has been written correctly.

5 Packaging and Reusing Code

5.1 Functions

Functions provide a way to reuse a block of code by giving it a name. The code inside the function can then be executed by calling the function name. Using functions makes code more understandable and easier to maintain.

- Complex problems are better understood when broken down into simpler steps. By putting the code for each step into a well-named function, the overall structure of the solution is easier to see and understand.
- Duplicating code in two or more places makes code longer and more difficult to maintain. This is because changes to the code need to be made each place the code is duplicated, leading to problems if the code is not kept in sync. Putting duplicated code into a function which is called wherever the code is needed avoids this problem.
- Variables defined and used inside a function only exist inside the function. They are said to be *local* variables, as opposed to *global* variables that can be accessed everywhere. Local variables allow the same variable name to be used safely in different parts of a program. This is because modifying the variable inside the function only changes its value inside the function, without affecting the rest of the program.

Functions can be defined with the option of passing in additional data to be used inside the function. The variables used to identify this additional data are the function *parameters*, and the particular values passed in when the function is called are the function *arguments*.

The **def** statement is used to define functions in Python, with the syntax:

def statement

```
def <name> (<parameters>):  
    """documentation string"""  
    <code>
```

- Functions take a list of required arguments, identified by position.
- Functions can take *keyword* arguments, identified by name. These can also be assigned default values in the function definition to use if no values are passed in.
- Functions can return one or more values using the **return** statement. Note that functions do not *have* to return a value - they could just perform some action instead. A function stops executing as soon as a return statement is encountered.

- An optional documentation string can be added at the start of the function (before the code) to describe what the function does. This string is usually enclosed in triple quotes.

The arguments to a function can be specified by position, keyword, or some combination of both. Some examples using just positional arguments are as follows.

| | |
|---|--|
| <pre>def square(x): return x**2 print(square(3)) 9</pre> | <p><i>The function exits as soon as the return statement is called. The value after the return statement is sent back to the code that called the function.</i></p> |
| <pre>def multiply(x, y): """Return the product xy""" return x*y print(multiply(3, 2)) 6</pre> | <p><i>Parameters are bound to the function arguments in the order given. The documentation string is placed after the colon and before the code.</i></p> |
| <pre>def minmax(data): return min(data), max(data) print(minmax([1, 3, 7, 2, 10])) (1, 10)</pre> | <p><i>Multiple values are returned as a tuple.</i></p> |

Using *keyword* arguments allows default values to be assigned. This is particularly useful when a function can be called with many different options, and avoids having to call functions with a long list of arguments.

- Keyword arguments are specified using `key=default` in place of a positional argument.
- Using keyword instead of positional arguments means we don't need to remember the order of arguments, and allows the defaults to be used most of the time.
- Positional and keyword arguments can be used in the same function, as long as the positional arguments come first.

| | |
|---|---|
| <pre>def close_enough(x, y, tolerance=.1) return abs(x - y) <= tolerance</pre> | <p><i>The tolerance argument is 0.1 by default.</i></p> |
|---|---|

| | |
|---|--|
| <code>close_enough(1, 1.05)</code> | <i>The default tolerance of 0.1 is used in this case.</i> |
| True | |
| <code>close_enough(1, 1.05, tolerance=.01)</code> | <i>The default tolerance is overridden by the value of 0.01.</i> |
| False | |

If the number of arguments is not known in advance, functions can be defined to take a variable number of positional arguments and/or a variable number of keyword arguments. We are unlikely to be using these options ourselves, although they occur frequently in the documentation for Matplotlib.

- The positional arguments are specified as `*args` and are available as a tuple. Individual positional arguments can then be accessed by indexing into the tuple by position.
- The keyword arguments are specified as `**kwargs` and are available as a dictionary. Individual keyword arguments can then be accessed by indexing into this dictionary by key.

An example using a variable number of positional arguments is given below.

| | |
|--|---|
| <pre>def poly(x, *coefficients): "Evaluate a polynomial at value x" total = 0 for n, a in enumerate(coefficients): total += a*(x**n) return total poly(10, 1, 2, 3)</pre> | <p><i>Defining a function to compute the polynomial</i></p> $a_0 + a_1x + a_2x^2 + \dots$ <p><i>The *coefficients argument can be a sequence of any length. The function call here computes</i></p> $1 + 2 \cdot 10 + 3 \cdot 10^2$ |
| 321 | |

We may on occasion need a simple function that will only be used in a single place, and not want to have to define and name a separate function for this purpose. In this case we can define an anonymous or *lambda* function just in the place where it is needed. The syntax for a lambda function is:

| |
|--|
| lambda statement |
| <code>lambda <arguments> : <code></code> |

The **lambda** statement returns an unnamed function which takes the arguments given before the colon, and returns the result of executing the code after the colon. Typical uses for lambda functions are where one function needs to be passed in as an argument to a different function.

```
ages = [21, 19, 98]
names = ["Bruce", "Sheila", "Adam"]
data = zip(ages, names)
sorted(data, key=lambda x : x[0])

[(19, "Sheila"), (21, "Bruce"), (98,
"Adam")]
```

The key argument to **sorted** lets us sort the data based on the first list. Using a lambda function avoids having to define and name a separate function for this simple task.

5.2 Modules

A module is a file containing Python definitions and statements. These allow us to use code created by other developers, and greatly extend what we can do with Python. Since many different modules are available, it is possible that the same names are used by different developers. We therefore need a way to identify *which* module a particular variable or function came from.

The **import** statement is used to make the variables and functions in a module available for use. We can either:

- Import everything from a module for immediate use.
- Import only certain named variables and functions from a module.
- Import everything from a module, but require that variable and function names be prefaced by either the module name or some alias.

```
from math import pi
print(pi)

3.14159265359
```

pi is now a variable name that we can use, but not the rest of the *math* module.

```
from math import *
print(e)

2.71828182846
```

Everything in the *math* module is now available.

```
import numpy
print(numpy.arcsin(1))

1.57079632679
```

Everything in *numpy* can be used, prefaced by "*numpy*".

```
import numpy as np
print(np.cos(0))

1.0
```

Everything in *numpy* can be used, prefaced by the alias "*np*".

If we want to know what a module contains, we can use the **dir** function. This returns a list of all the variable and function names in the module.


```
import math
print(dir(math))
```

```
["__doc__", "__name__", "__package__",
 "acos", "acosh", "asin", "asinh",
 "atan", "atan2", "atanh", "ceil",
 "copysign", "cos", "cosh", "degrees",
 "e", "erf", "erfc", "exp", "expm1",
 "fabs", "factorial", "floor", "fmod",
 "frexp", "fsum", "gamma", "hypot",
 "isinf", "isnan", "ldexp", "lgamma",
 "log", "log10", "log1p", "modf",
 "pi", "pow", "radians", "sin",
 "sinh", "sqrt", "tan", "tanh",
 "trunc"]
```

The **math** module contains all the standard mathematical functions, as well as the constants "e" and "pi".

5.3 Comprehensions

Often we want to create a container by modifying and filtering the elements of some other sequence. Comprehensions provide an elegant way to do this, similar to mathematical set-builder notation. For list comprehensions, the syntax is:

List comprehension

```
[<expression> for <variables> in <iterable> if <condition>]
```

The code in \langle expression \rangle is evaluated for each item in the \langle container \rangle , and the result becomes an element of the new list. The \langle condition \rangle does not have to be present but, if it is, only elements which satisfy the condition become incorporated into the new list.

```
[i**2 for i in range(5)]
```

```
[0, 1, 4, 9, 16]
```

*i**2 is evaluated for every item i in the list.*

```
[d for d in range(1,7) if 6 % d == 0]
```

```
[1, 2, 3, 6]
```

Divisors of 6 - only elements passing the test 6 % d == 0 are included.

We can also use a dictionary comprehension to create a dictionary without needing to repeatedly add key:value pairs.

Dictionary comprehension

```
{<key>:<value> for <variables> in <iterable> if <condition>}
```

The following example creates a dictionary from a sequence of integers.

```
{i:i**2 for i in range(4)}
{0:0, 1:1, 2:4, 3:9}
```

Create a dictionary from a sequence of integers. Note the key:value pairs and surrounding curly brackets.

5.4 Generator Expressions

We often want to iterate over the elements of some sequence, but don't need to save the sequence. In these cases, a list comprehension is an unnecessary overhead since the list is created and saved in memory for a single use. A more efficient alternative is to use a "generator expression".

- Generator expressions are used like lists inside **for** loops.
- They don't create the entire sequence before it is used, so memory doesn't need to be allocated to store all the elements.
- Instead, the current element is saved, and the next one is produced only when requested by the loop.

Generator expressions have almost the same syntax as a list comprehension, but use parentheses instead of square brackets.

Generator expression

```
(<expression> for <variables> in <iterable> if <condition>)
```

An example is given below.

```
squares = (i**2 for i in range(1, 4))
for s in squares:
    print(s)
1
4
9
```

The elements of squares are generated as needed, not generated and saved in advance.

5.5 Comments

Comments are text that is included in the code but not executed. They are used to document and explain what the code is doing. Python allows two forms of comment.

- A hash symbol `#` means that the rest of the line is a comment, and is not to be executed.

- A documentation string is surrounded by triple quotes `"""`. Everything inside the quotes is ignored.

```
# This is a single-line comment
```

No need to comment a comment.

```
"""This is a documentation string."""
```

III Numerical Computing

6 NumPy

NumPy (Numerical Python) is the fundamental package for scientific computing with Python. It defines a new kind of container - the ndarray (usually just referred to as an array) - that supports fast and efficient computation. NumPy also defines the basic routines for accessing and manipulating these arrays.

Arrays have the following properties (among others):

- A *shape*, which is a tuple of integers. The number of integers is the number of dimensions in the array, and the integers specify the size of each dimension.
- A *dtype* (data-type), which specifies the type of the objects stored in the array.

In NumPy, the dimensions of an array are referred to as *axes*. An example of an array with dtype int and shape ((4, 5)) is shown below. The first axis has four elements, each of which is a one-dimensional array with 5 elements.

```
[[ 0  1  2  3  4 ]
 [ 5  6  7  8  9 ]
 [10 11 12 13 14 ]
 [15 16 17 18 19 ]]
```

The main differences between NumPy arrays and Python lists are:

- The objects in a NumPy array must all be of the same type - booleans, integers, floats, complex numbers or strings.
- The size of an array is fixed at creation, and can't be changed later.
- Arrays can be multi-dimensional.
- Mathematical operations can be applied directly to arrays. When this is done they are applied elementwise to the array, generating another array as output. This is *much* faster than iterating over a list.
- Indexing for arrays is more powerful than that for lists, and includes indexing using integer and boolean arrays.
- Slicing an array produces a view of the original array, not a copy. Modifying this view will change the original array.

NumPy is well documented online, with a [standard tutorial](#) and [good introductory tutorial](#) available.

6.1 Array Creation

NumPy arrays can be created:

- From a list. The elements of the list need to all be of the same type, or of a kind that can all be cast to the same type. For example, a list consisting of both integers and floats will generate an array of floats, since the integers can all be converted to floats.
- According to a given shape. The array will be initialized differently depending on the function used.
- From another array. The new array will be of the same shape as the existing array, and could either be a copy, or initialized with some other values.
- As a result of an operation on other arrays. The standard mathematical operators can all be applied directly to arrays. The result is an array of the same shape where the operation has been performed separately on corresponding elements.

The following functions are the main ones we use in this class.

| | |
|-------------------------|--|
| <code>array</code> | Create an array from a list. |
| <code>linspace</code> | Return an array of evenly spaced numbers over a specified interval. |
| <code>arange</code> | Return an array of evenly spaced integers within a given interval. |
| <code>empty</code> | Return an a new array of a given shape and type, without initializing entries. |
| <code>zeros</code> | Return an a new array of a given shape and type, filled with zeros. |
| <code>ones</code> | Return an a new array of a given shape and type, filled with ones. |
| <code>empty_like</code> | Return a new array with the same shape and type as a given array. |
| <code>zeros_like</code> | Return an array of zeros with the same shape and type as a given array. |
| <code>ones_like</code> | Return an array of ones with the same shape and type as a given array. |
| <code>copy</code> | Return an array copy of the given object. |
| <code>meshgrid</code> | Returns a pair of 2D x and y grid arrays from 1D x and y coordinate arrays. |

These functions are illustrated below. The "np" alias is standard practice when using NumPy.

```
import numpy as np
x = np.array([1, 2, 3])
print(x)
```

```
[1 2 3]
```

array(object) creates an array from a list - note that arrays are printed without commas.

| | |
|---|--|
| <pre>x = np.array([1, 2, 3], dtype=float) print(x)</pre> <hr/> <pre>[1. 2. 3.]</pre> | <p>array(<i>object, dtype</i>) creates an array of type <i>dtype</i> - the integers are now cast to floats.</p> |
| <pre>x = np.linspace(0, 1, 6) print(x)</pre> <hr/> <pre>[0. 0.2 0.4 0.6 0.8 1.]</pre> | <p>linspace(<i>start, stop, num</i>) returns <i>num</i> equally spaced points, including endpoints.</p> |
| <pre>x = np.arange(5) print(x)</pre> <hr/> <pre>[0 1 2 3 4]</pre> | <p>arange returns an array of evenly spaced values within a given interval.</p> |

The functions **empty**, **zeros** and **ones** all take a *shape* argument and create an array of that shape, initialized as appropriate.

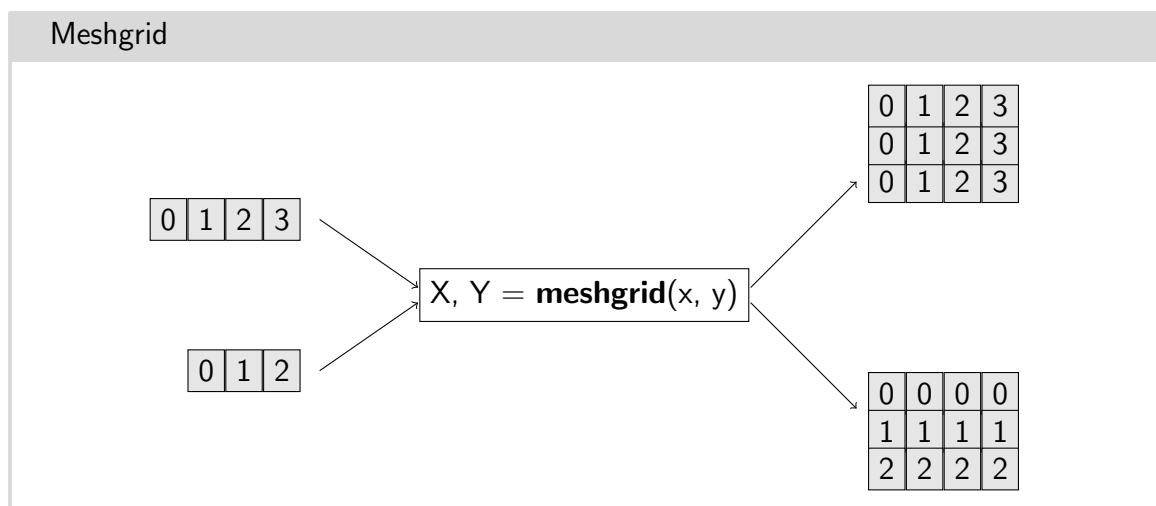
| | |
|---|---|
| <pre>x = np.empty((3, 2)) print(x)</pre> <hr/> <pre>[[6.93946206e-310 6.93946206e-310] [6.36598737e-314 6.36598737e-314] [6.36598737e-314 0.00000000e+000]]</pre> | <p>empty(<i>shape</i>) returns an array of shape <i>shape</i>, initially filled with garbage.</p> |
| <pre>x = np.zeros((2, 3)) print(x)</pre> <hr/> <pre>[[0. 0. 0.] [0. 0. 0.]]</pre> | <p>zeros(<i>shape</i>) returns an array of shape <i>shape</i> filled with zeros - note the default type is float.</p> |
| <pre>x = np.ones((2, 3), dtype=int) print(x)</pre> <hr/> <pre>[[1 1 1] [1 1 1]]</pre> | <p>ones(<i>shape, dtype</i>) returns an array of shape <i>shape</i> filled with ones - using <i>dtype=int</i> casts the elements to type <i>int</i>.</p> |

Arrays can be created directly from other arrays using **empty_like**, **zeros_like**, **ones_like** and **copy**.

| | |
|---|---|
| <pre>x = np.arange(3, dtype=float) print(x)</pre> <hr/> <pre>[0. 1. 2.]</pre> | <p>Create an array of floats using arange.</p> |
|---|---|

| | |
|---|---|
| <pre>y = np.empty_like(x) print(y)</pre> <hr/> <pre>[0.00000000e+000 6.51913678e+091 6.95022185e-310]</pre> | <i>y has the same shape as x, but is initially filled with garbage.</i> |
| <pre>y = np.zeros_like(x) print(y)</pre> <hr/> <pre>[0. 0. 0.]</pre> | <i>y has the same shape as x, but is initialized with zeros.</i> |
| <pre>y = np.ones_like(x) print(y)</pre> <hr/> <pre>[1. 1. 1.]</pre> | <i>y has the same shape as x, but is initialized with ones.</i> |
| <pre>y = np.copy(x) print(y)</pre> <hr/> <pre>[0. 1. 2.]</pre> | <i>y is a copy of x - changing y will not change x.</i> |

The function **meshgrid**(*x*, *y*) creates two-dimensional arrays from one-dimensional *x*- and *y*-coordinate axes. One array contains the *x*-coordinates of all the points in the *xy*-plane defined by these axes, and the other contains the *y*-coordinates. An example is shown below.



| | |
|---|---|
| <pre>x = np.arange(4) y = np.arange(3) X, Y = np.meshgrid(x, y)</pre> <hr/> | meshgrid creates 2D <i>x</i> - and <i>y</i> - coordinate arrays from 1D <i>x</i> - and <i>y</i> - coordinate arrays. |
|---|---|

| | |
|---|---|
| <pre>print(X) [[0 1 2 3] [0 1 2 3] [0 1 2 3]]</pre> | <i>X is a 2D array containing just the x-coordinates of points in the xy plane.</i> |
| <pre>print(Y) [[0 0 0 0] [1 1 1 1] [2 2 2 2]]</pre> | <i>Y is a 2D array containing just the y-coordinates of points in the xy plane.</i> |

The function **meshgrid** is often used when we want to apply a function of two variables to points in the x-y plane. The following example uses the arrays *X* and *Y* above.

| | |
|---|---|
| <pre>def distance(x, y): return np.round(sqrt(x**2 + y**2), 3) print(distance(X, Y))</pre> | <i>The function distance finds the distance of a point (x, y) from the origin, rounded to 3 decimal places by the round function.</i> |
| <pre>[[0. 1. 2. 3.] [1. 1.414 2.236 3.162] [2. 2.236 2.828 3.606]]</pre> | |

6.2 Array Properties

The properties of an array can be accessed as follows.

| | |
|--|--|
| <pre>import numpy as np x = np.arange(6) type(x)</pre> | <i>x is of type <code>numpy.ndarray</code>.</i> |
| <pre>numpy.ndarray</pre> | |
| <pre>x.dtype</pre> | <i>dtype returns the element type - a 64-bit integer.</i> |
| <pre>dtype("int64")</pre> | |
| <pre>x.shape</pre> | <i>x is a 1-dimensional array with 6 elements in the first axis.</i> |
| <pre>(6,)</pre> | |

It is possible to create an array from the elements of an existing array, but with the properties changed.

- The number of dimensions and size of each dimension can be changed using **reshape** (as long as the total number of elements is unchanged).
- The dtype can be changed using **astype**.

| | |
|--|---|
| <pre>x = np.arange(6).reshape((2, 3)) print(x)</pre> <hr/> <pre>[[0 1 2] [3 4 5]]</pre> | <p>reshape creates a view of an array with the same number of elements, but a different shape.</p> |
| <pre>y = x.astype(float) print(y)</pre> <hr/> <pre>[[0. 1. 2.] [3. 4. 5.]]</pre> | <p>astype casts the integers in x to floats in y. This creates a new array - modifying it will not alter the original.</p> |

6.3 Array Operations

Array arithmetic is done on an element-by-element basis. Operations are applied to every element of an array, with each result becoming an element in a new array.

| | |
|--|--|
| <pre>import numpy as np x = np.arange(4) print(x)</pre> <hr/> <pre>[0 1 2 3]</pre> | <p>Create an array of consecutive integers using arange.</p> |
| <pre>print(x + 1)</pre> <hr/> <pre>[1 2 3 4]</pre> | <p>1 is added to every element of the array x.</p> |
| <pre>print(x * 2)</pre> <hr/> <pre>[0 2 4 6]</pre> | <p>Every element of the array x is multiplied by 2.</p> |
| <pre>print(x ** 2)</pre> <hr/> <pre>[0 1 4 9]</pre> | <p>Every element of the array x is squared.</p> |
| <pre>y = array([3, 2, 5, 1])</pre> <hr/> | <p>Create a second array</p> |

| | |
|--|---|
| <pre>print(x) print(y) print(x + y)</pre> <hr/> <pre>[0 1 2 3] [3 2 5 1] [3 3 7 4]</pre> | <p>The elements of x are added to the corresponding elements of y on an element-by-element basis.</p> |
| <pre>print(x**y)</pre> <hr/> <pre>[0 1 32 3]</pre> | <p>Exponentiation is done using corresponding elements of the arrays x and y.</p> |

Comparison operators and other Boolean expressions are also applied on an element-by-element basis. The Boolean expression is applied to every element of the array, with each result becoming an element in a new boolean array.

| | |
|---|---|
| <pre>x = np.arange(5) print(x) print(x % 2 == 0)</pre> <hr/> <pre>[0 1 2 3 4] [True False True False True]</pre> | <p>The Boolean expression is evaluated for each element separately, resulting in an array of booleans.</p> |
| <pre>x = np.arange(4) y = np.array([3, 2, 5, 1]) print(x) print(y) print(x < y)</pre> <hr/> <pre>[0 1 2 3] [3 2 5 1] [True True True False]</pre> | <p>The comparison is done on an elementwise basis between elements of arrays x and y. The result is an array of booleans.</p> |

NumPy contains vectorized versions of all the basic mathematical functions. Note that they need to be imported before we can use them. Some examples are given below.

| | |
|--|--|
| <pre>x = np.arange(3) print(x)</pre> <hr/> <pre>[0 1 2]</pre> | <p>Create an array.</p> |
| <pre>print(np.sin(x))</pre> <hr/> <pre>[0. 0.84147098 0.90929743]</pre> | <p>sin is applied to each element, to create a new array.</p> |

| | |
|---|--|
| <pre>print(np.exp(x))</pre> <hr/> <pre>[1. 2.71828183 7.3890561]</pre> | <p>exp is the exponential operator.</p> |
| <pre>x = np.random.randint(5, size=(2,3))</pre> <pre>print(x)</pre> <hr/> <pre>[[1 4 3] [2 2 3]]</pre> | <p>random.randint returns an array of a given size filled with randomly selected integers from a given range.</p> |
| <pre>print(x.min(), x.max())</pre> <hr/> <pre>1 4</pre> | <p>min and max calculate the minimum and maximum values across the entire array.</p> |
| <pre>print(x.min(axis=0))</pre> <pre>print(x.min(axis=1))</pre> <hr/> <pre>[1 2 3] [1 2]</pre> | <p>The axis argument finds each minimum along a given axis. The resulting array is the shape of the original array, but with the given axis removed.</p> |
| <pre>print(x.sum())</pre> <hr/> <pre>15</pre> | <p>sum sums all the elements of an array.</p> |
| <pre>print(x.sum(axis=0))</pre> <hr/> <pre>[3 6 6]</pre> | <p>Providing the axis argument sums along the given axis.</p> |

6.4 Array Indexing and Slicing

Arrays can be indexed and sliced using square brackets in the same way as lists.

- An index or start:stop:step filter needs to be provided for each axis, separated by commas.
- Indexing is zero-based, as it is with lists.
- Slicing an array produces a view of the original array, not a copy. Modifying this view will change the original array.

| | |
|--|---|
| <pre>import numpy as np x = np.arange(20).reshape((4, 5)) print(x)</pre> <pre>[[0 1 2 3 4] [5 6 7 8 9] [10 11 12 13 14] [15 16 17 18 19]]</pre> | <p>reshape provides a fast way to create a 2D array from a 1D array.</p> |
| <pre>print(x[1,2])</pre> <pre>7</pre> | <p>Indexing is done into each axis in order - row 1, column 2.</p> |
| <pre>print(x[1])</pre> <pre>[5 6 7 8 9]</pre> | <p>Row 1 can be selected using an integer.</p> |
| <pre>print(x[:,1])</pre> <pre>[1 6 11 16]</pre> | <p>Slicing can be used to select the first element of every axis (i.e. column 1).</p> |
| <pre>print(x[1:3, 1:4])</pre> <pre>[[6 7 8] [11 12 13]]</pre> | <p>Slice rows 1 and 2 using 1:3, then slice columns 1, 2 and 3 using 1:4.</p> |

6.5 Indexing with Integer Arrays

Although slicing can be used to return a subset of elements from an array, its use is restricted to either selecting consecutive elements, or to selecting elements that are separated by the same fixed amount. We sometimes want to select instead some *arbitrary* subset of an array, possibly even selecting the same element more than once. Integer array indexing provides a way to do this.

- The index is a NumPy array of integers.
- Each integer in the index array selects a corresponding element from the target array.
- The result is an array where the first axes are the same shape as the index array.

Using an integer array as an index

| | | | | | | | | | |
|----------|---|----|---|---|----|----|----|----|----|
| x | 0 | 1 | 4 | 9 | 16 | 25 | 36 | 49 | 64 |
| index | 2 | 5 | | | | | | | |
| x[index] | 4 | 25 | | | | | | | |

```
import numpy as np
x = np.arange(9)**2
print(x)
```

```
[ 0  1  4  9 16 25 36 49 64]
```

First create the array using **arange**, then square each element.

```
index = np.array([2, 5])
print(x[index])
```

```
[4 25]
```

An array is returned containing elements from the first array, selected according to the integers in the second array.

The index can be an integer array of any shape, and can include duplicates as shown below.

```
index = np.array([[2, 3], [7, 2]])
print(index)
```

```
[[2 3]
 [7 2]]
```

The indexing array contains integers that are used to index into the target array. Note that the same elements (2, in this case) can be selected more than once.

```
print(x[index])
```

```
[[ 4  9]
 [49  4]]
```

When indexing a one-dimensional array using an integer array, the returned array has the same shape as the indexing array.

The target array can also be a multidimensional array. In this case, the elements selected by the index array will be the array elements of the target.

```

colors = np.array([[1., 0., 0.],
                  [0., 1., 0.],
                  [0., 0., 1.]])
index = np.array([1, 0, 2, 1, 1, 0])
print(colors[index])

```

```

[[0.  1.  0.]
 [1.  0.  0.]
 [0.  0.  1.]
 [0.  1.  0.]
 [0.  1.  0.]
 [1.  0.  0.]]

```

The colors array is a two-dimensional array, so the elements of colors are one-dimensional arrays (the rows).

The colors array can be thought of as a "lookup table" - integer indexing looks up the elements of this table, and uses the values found to construct the result.

6.6 Indexing with Boolean Arrays

We can also index using arrays of booleans. In this case, the indexing array acts like a *mask*, letting through only the elements in the target array that correspond to True values in the indexing array.

Using a boolean array as a mask

| | | | | |
|---------|------|------|-------|------|
| x | 0 | 1 | 2 | 3 |
| mask | True | True | False | True |
| x[mask] | 0 | 1 | 3 | |

```

import numpy as np
x = np.arange(4)
mask = np.array([True, True, False, True])
print(x)
print(x[mask])

```

```

[0 1 2 3]
[0 1 3]

```

Only the elements with a matching True in the mask are selected.

A common use of this technique is to select the elements of an array that satisfy some condition. This can be done by first applying the condition to the array to generate an array of booleans, then using the resulting array as an index. The result is that only the elements for which the condition holds true are selected.

Selecting array elements that satisfy a condition

| | | | | | | | | | | | | | | | | | | | | |
|---------------|---|---|---|---|----|----|----|---|---|---|----|----|----|----|----|----|----|----|----|----|
| x | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |
| x % 3 == 0 | T | F | F | T | F | F | T | F | F | T | F | F | T | F | F | T | F | F | T | F |
| x[x % 3 == 0] | 0 | 3 | 6 | 9 | 12 | 15 | 18 | | | | | | | | | | | | | |

```
x = np.arange(20)
index3 = (x % 3 == 0)
index5 = (x % 5 == 0)
```

index3 and index5 are boolean arrays containing True elements for the integers that are divisible by 3 and 5 respectively.

```
print(x[index3])
[ 0  3  6  9 12 15 18]
```

Select just the elements of x that are divisible by 3.

```
print(x[index5])
[ 0  5 10 15]
```

Select just the elements of x that are divisible by 5.

```
print(x[np.logical_or(index3, index5)])
[ 0  3  5  6  9 10 12 15 18]
```

*The function **logical_or** performs an elementwise "or". The result is the integers divisible by either 3 or 5.*



The logical operators **not**, **or** and **and** do not get applied elementwise when applied to NumPy arrays. The functions **logical_not**, **logical_or** and **logical_and** need to be used instead.

7 Matplotlib

Matplotlib is a 2D plotting library for Python. It can be used to generate graphs, histograms, bar charts, contour plots, polar plots, scatter plots, and many other kinds of mathematical graphics. The "pyplot" interface provides a MATLAB-like interface for simple plotting, and is the main one we will be using in class. The [online reference](#) provides a full description of the available functions. A [good tutorial](#) is also available online.

The following commands are the main ones used for creating and formatting graphs.

| | |
|----------------------------|---|
| <code>plot</code> | Plot lines and/or markers. |
| <code>scatter</code> | Scatter plot of x vs y. |
| <code>show</code> | Display a figure. |
| <code>title</code> | Set a title for the graph. |
| <code>xlabel/ylabel</code> | Set labels for the x and y axes. |
| <code>xlim/ylim</code> | Get or set the range of x and y values to be displayed. |
| <code>xticks/yticks</code> | Get or set the locations and labels for the tick marks on the x and y axes. |
| <code>subplot</code> | Plot multiple graphs in one figure. |
| <code>figure</code> | Create a new figure. |
| <code>fill_between</code> | Fill the area between two curves. |
| <code>legend</code> | Put a legend on the graph. |

Colors, line styles, and marker styles can all be set to create customized graphs. These are usually specified as strings, with the most frequently used options as follows.

| Style options | | Colors | |
|---------------|----------------|--------|---------|
| "-" | solid line | "b" | blue |
| "- -" | dashed line | "g" | green |
| "." | point marker | "r" | red |
| "o" | circle marker | "c" | cyan |
| "s" | square marker | "m" | magenta |
| "+" | plus marker | "y" | yellow |
| "x" | x marker | "k" | black |
| "D" | diamond marker | "w" | white |

7.1 Basic Plotting

Functions can be graphed using a call to `plot(x, y)`, followed by a call to `show`. Note that:

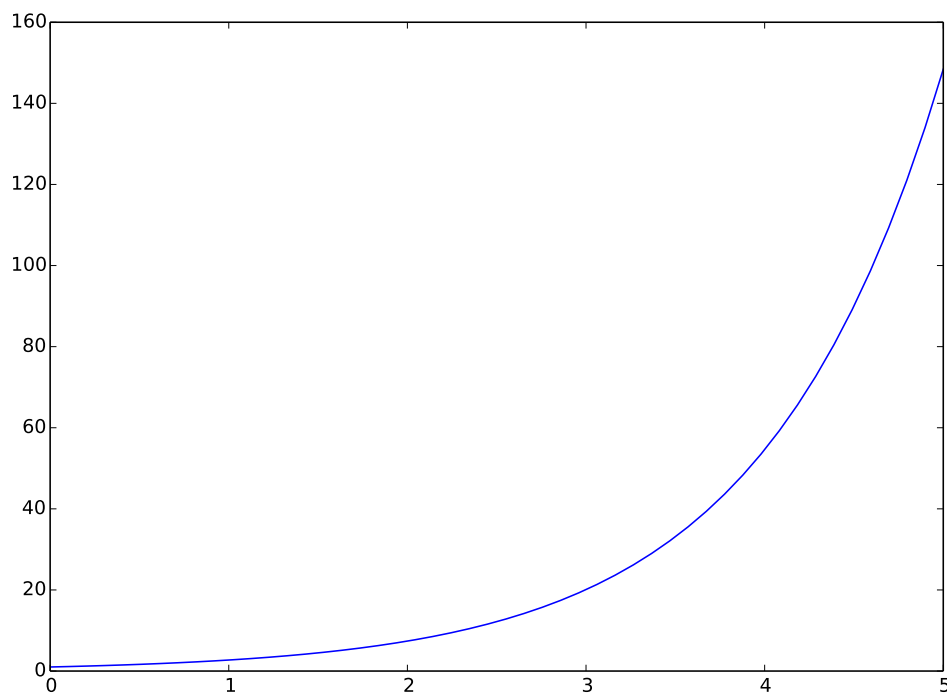
- The `x` parameter contains the x-coordinates of the points to plot, and the `y` parameter contains the y-coordinates.
- We need to import the required NumPy and Matplotlib functions for array manipulation and plotting.
- If using Jupyter Notebook, calling the IPython magic `%matplotlib inline` calls `show` automatically when the cell is executed.
- The default plotting behavior is to connect the points with a blue line.

The following example plots the exponential function in the range $[0, 5]$.

Plotting an Exponential Function

```
import numpy as np          # NumPy functions available with "np." prefix
import matplotlib.pyplot as plt # Import plotting functions from Matplotlib

x = np.linspace(0, 5)      # Create array of equally spaced values
plt.plot(x, np.exp(x))     # Plot the exponential function
plt.show()                 # Finally, show the figure
```



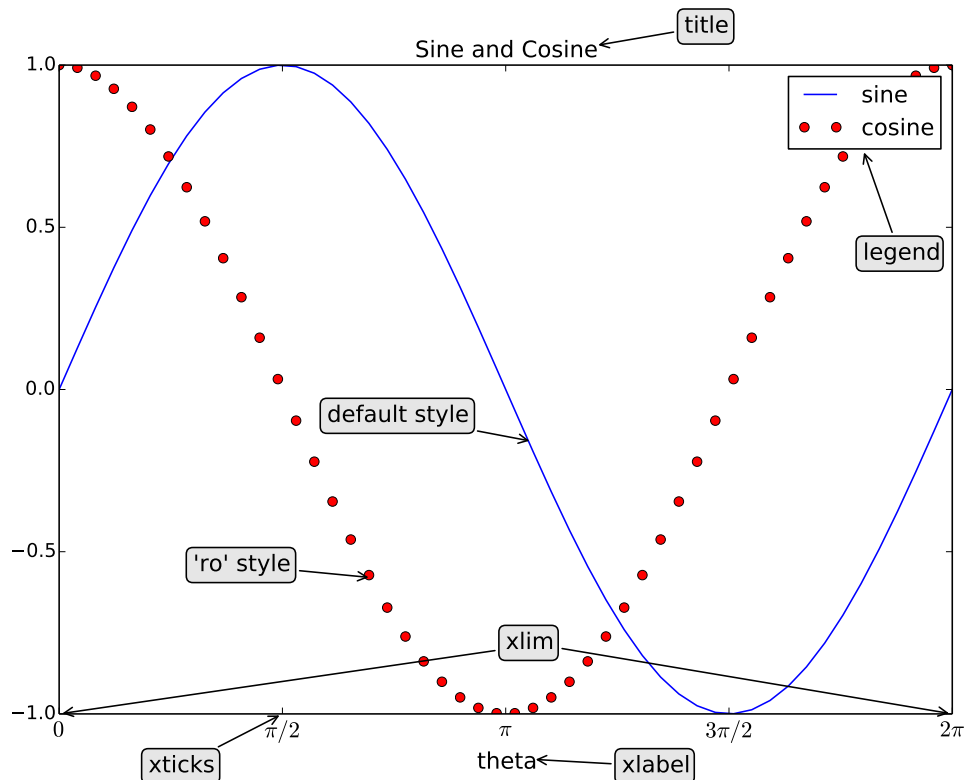
7.2 A More Complex Plotting Example

A range of options are available for customizing plots. These are illustrated in the example below, which plots a sine and cosine curve on the same graph. Note that:

- The third argument to plot can be used to set colors, line types and marker types.
- Multiple plots can be drawn on one figure, followed by a single call to **show**.

Plotting Two Graphs on a Single Figure

```
import numpy as np                                # Imports linspace, sin, cos
import matplotlib.pyplot as plt                  # Import plotting functions
x = np.linspace(0, 2*np.pi, 50)                # Plot 50 points on the x-axis
plt.figure(figsize=(10,7))                      # Set the size of the figure
plt.plot(x, np.sin(x), label='sine')            # Default style is a blue line
plt.scatter(x, np.cos(x), c='r', label='cosine') # Use 'r' for red circles
plt.xlabel('theta')                             # Label the x-axis
plt.xlim(0, 2*np.pi)                          # Limit x-axis to this range
ticks = [i*np.pi/2 for i in range(5)]         # Locations of ticks on x-axis
labels = [r'$0$', r'$\pi/2$', r'$\pi$',        # Labels for the x-axis ticks
          r'$3\pi/2$', r'$2\pi$']              # - these are LaTeX strings
plt.xticks(ticks, labels, size='large')         # Add the x-ticks and labels
plt.title('Sine and Cosine')                   # Add a title
plt.legend()                                    # Legend uses the plot labels
plt.show()                                     # Finally, show the figure
```



7.3 Bar Plots

The function `bar` is used to create bar plots.

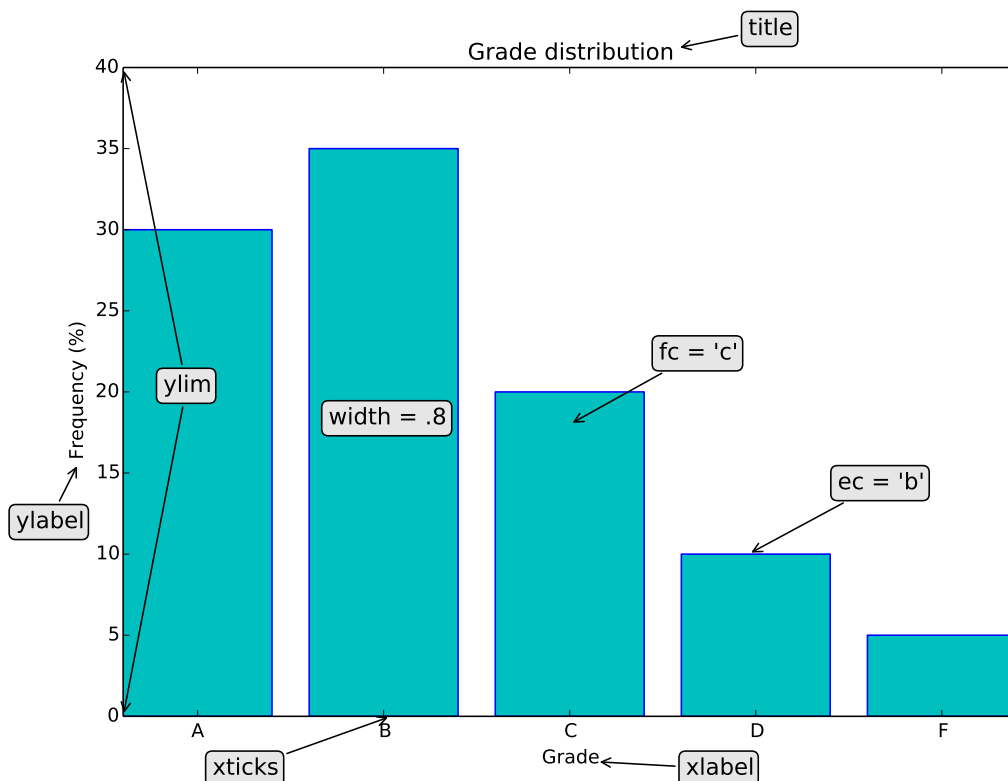
- Bars are described by their height, width, and position of the left and bottom edges.
- The `width` argument can be used to make bars thinner or thicker.
- The face color and edge color of the bars can be specified independently.

The following example shows a bar plot with the face color set to "c" (cyan) and edge color set to "b" (blue). Labels are positioned at the centers of the bars.

Bar Plot

```
import matplotlib.pyplot as plt           # Import plotting functions

grades = ['A', 'B', 'C', 'D', 'F']      # Used to label the bars
freqs = [30, 35, 20, 10, 5]            # Bar heights are frequencies
width = 0.8                             # Relative width of each bar
ticks = [width/2 + i for i in range(5)] # Ticks in center of the bars
plt.bar(range(5), freqs, fc='c', ec='b') # fc/ec are face/edge colors
plt.xticks(ticks, grades)               # Place labels for the bars
plt.ylim(0, 40)                         # Set the space at the top
plt.title('Grade distribution')          # Add a title
plt.xlabel('Grade')                     # Add a label for the x-axis
plt.ylabel('Frequency (%)')             # Add a label for the y-axis
plt.show()                              # Finally, show the figure
```



7.4 Polar Plots

The function **polar** is used to create polar plots. These plot radius against angle in polar coordinates.

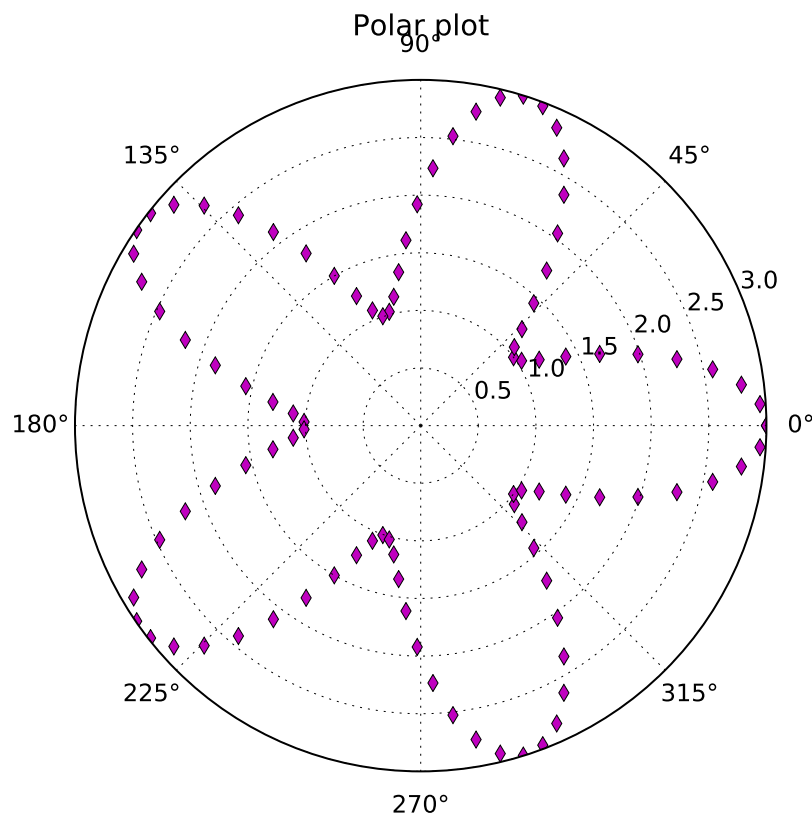
- The first argument to **polar** is an array of angles, and the second argument the corresponding radii.
- Colors, line types and marker types are specified in the same way as **plot**.
- Polar can be called multiple times, followed by a single call to **show**.

The following example shows a polar plot with the marker style to "d" (diamond) and the color set to "m" (magenta).

Plotting in Polar Coordinates

```
import numpy as np          # Import linspace, cos, pi
import matplotlib.pyplot as plt # Import plotting functions

theta = np.linspace(0, 2*np.pi, 100) # Create array of equally spaced values
r = 2 + np.cos(5*theta)              # Generate radius as a function of angle
plt.polar(theta, r, marker='d', ls='None', color='m') # Diamond marker 'd'
plt.title('Polar plot')              # Add the title
plt.show()                           # Finally, show the figure
```



7.5 Histograms

The function **hist** is used to plot histograms. These group numerical data into "bins", usually of equal width, in order to show how the data is distributed.

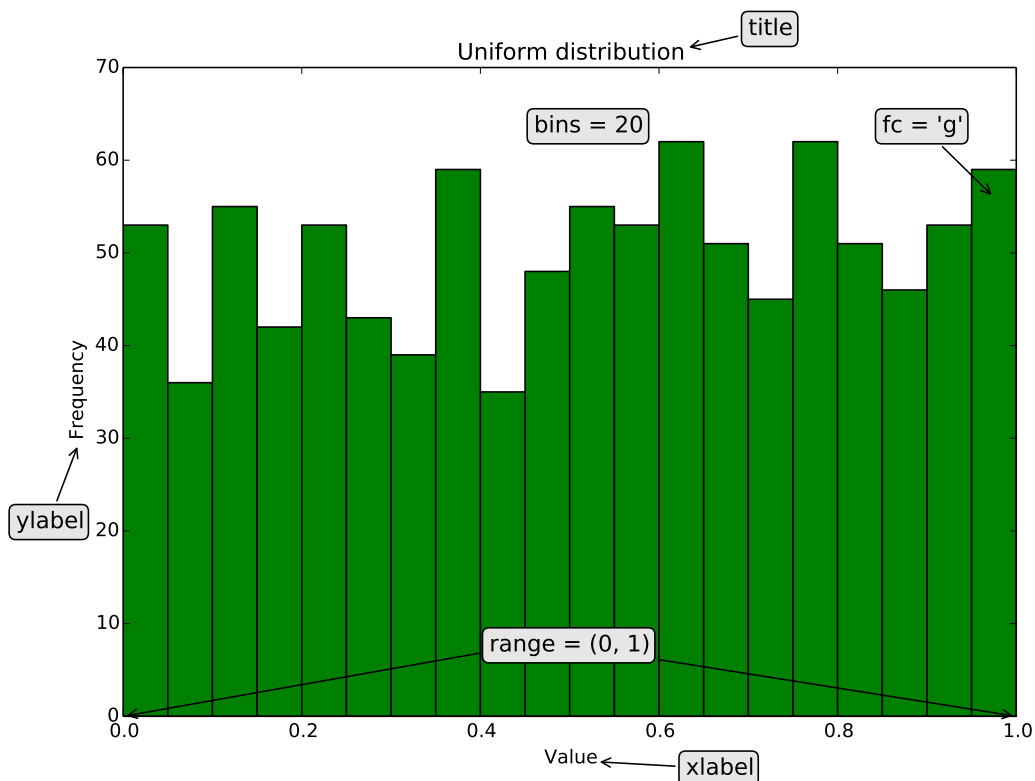
- Each bin covers a range of values, with the height of each bin indicating the number of points falling in that range.
- The first argument is an array or sequence of arrays.
- The *bins* argument specifies the number of bins to use.
- The *range* argument specifies the range of values to include.

The following example plots a histogram of 1000 samples drawn from a uniform probability distribution over $[0, 1)$.

Plotting a Histogram

```
import numpy as np                # Make random.rand available
import matplotlib.pyplot as plt   # Import plotting functions

x = np.random.rand(1000)         # 1000 random values in [0, 1)
plt.hist(x, bins=20, range=(0,1), fc='g') # Create histogram with 20 bins
plt.title('Uniform distribution')      # Add a title
plt.xlabel('Value')                 # Add a label for the x-axis
plt.ylabel('Frequency')             # Add a label for the y-axis
plt.show()                         # Finally, show the figure
```



7.6 Pie Charts

The function **pie** is used to create pie charts. These are a type of graph in which a circle is divided into wedges that each represent a proportion of the whole.

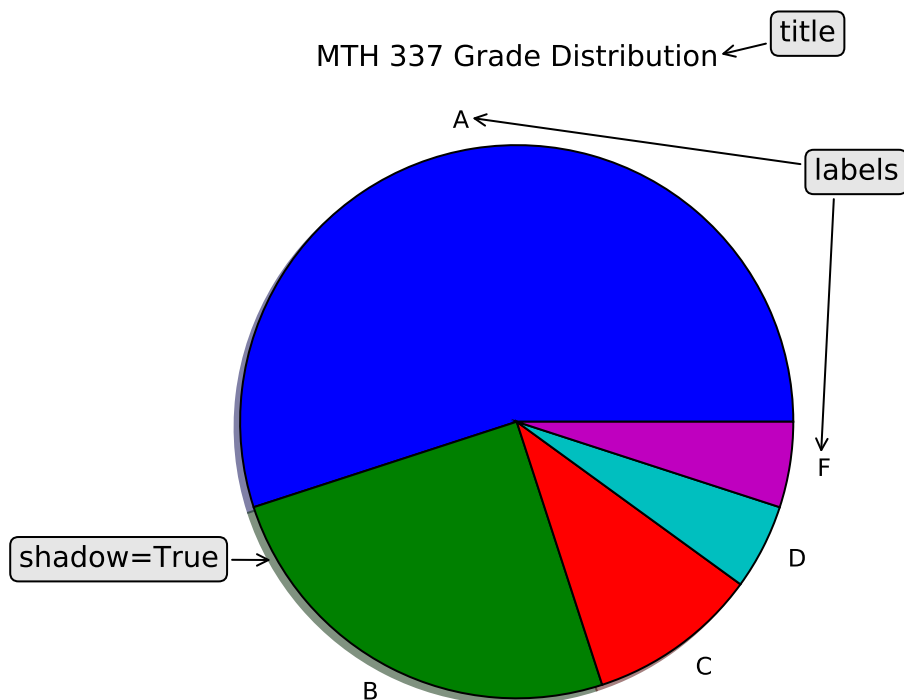
- The first argument to **pie** is a sequence of values used for the wedge sizes.
- The *labels* argument is a sequence of strings providing the labels for each wedge.
- The *shadow* argument is a boolean specifying whether to draw a shadow beneath the pie.

The following example shows a pie chart with *shadow* set to `True` .

Plotting a Pie Chart

```
import matplotlib.pyplot as plt          # Import plotting functions

percentages = [55, 25, 10, 5, 5]        # Wedge sizes
labels = ['A', 'B', 'C', 'D', 'F']      # List of labels for the wedges
plt.axes(aspect=1)                       # Aspect ratio = 1 for a true circle
plt.pie(percentages, labels=labels, shadow=True)
plt.title('MTH 337 Grade Distribution') # Add a title
plt.show()                                # Finally, show the figure
```



7.1 Contour Plots

The functions `contour` and `contourf` are used for contour plots and filled contour plots respectively. These are projections of a graph surface onto a plane, with the contours showing the level curves of the graph.

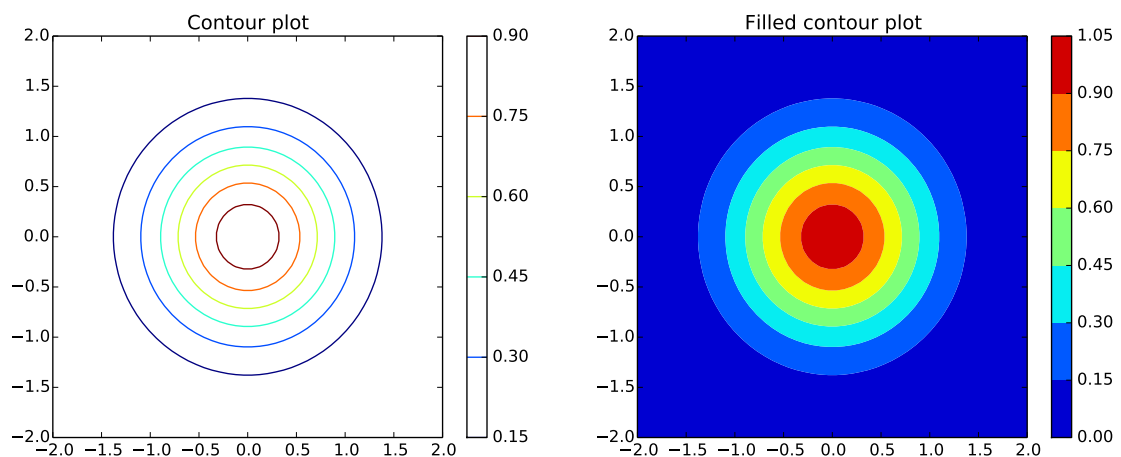
- The first two arguments are one dimensional arrays representing the x- and y-coordinates of the points to plot.
- The third coordinate is a two dimensional array representing the z-coordinates.
- Contour levels are automatically set, although they can be customized.
- A colorbar can be added to display the level curves.

The following examples are of a filled and unfilled contour plot of the two-dimensional Gaussian function, $f(x, y) = e^{-(x^2+y^2)}$.

Filled and Unfilled Contour Plots

```
import numpy as np                # Imports linspace, meshgrid, exp
import matplotlib.pyplot as plt   # Import plotting functions

x = np.linspace(-2,2)            # Locations of x-coordinates
y = np.linspace(-2,2)            # Locations of y-coordinates
XX, YY = np.meshgrid(x, y)       # meshgrid returns two 2D arrays
z = np.exp(-(XX**2 + YY**2))     # z is a 2D Gaussian
plt.figure(figsize=(14,5))       # Set the figure dimensions
plt.subplot('121')               # First subplot, 1 row, 2 columns
plt.contour(x, y, z)             # Contour plot
plt.title('Contour plot')        # Title added to first subplot
plt.colorbar()                   # Color bar added to first subplot
plt.subplot('122')               # Second subplot
plt.contourf(x, y, z)            # Filled contour plot
plt.title('Filled contour plot') # Title added to second subplot
plt.colorbar()                   # Color bar added to second subplot
plt.show()                       # Finally, show the figure
```



7.8 Slope Fields

The function **quiver** plots a two dimensional field of arrows, also known as a slope field.

- If four or more arguments are present, the first two arguments are one dimensional arrays representing the x- and y-coordinates of the arrows.
- The next two required arguments are two dimensional arrays representing the x- and y-components of the arrow vectors.
- An optional argument can be used to change the individual colors of the arrows. Keyword arguments allow the shape and appearance of the arrows to be customized.

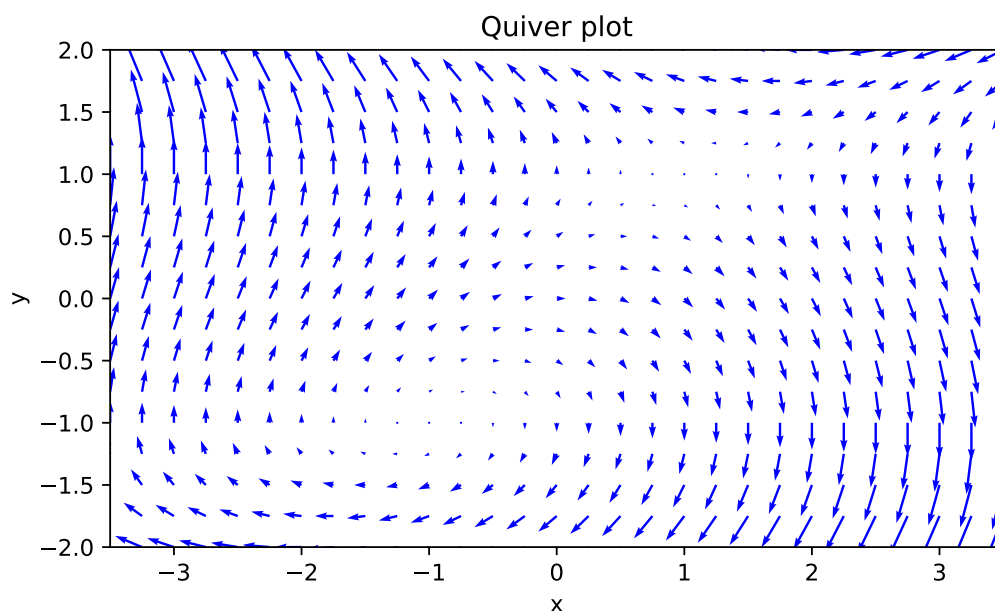
The following example is of a quiver plot of the system of ordinary differential equations,

$$\frac{dx}{dt} = 1 - y^2, \quad \frac{dy}{dt} = y - x$$

Plotting a Slope Field

```
import numpy as np          # Imports linspace, meshgrid
import matplotlib.pyplot as plt # Import plotting functions

x = np.linspace(-3.5, 3.5, 29) # Locations of x-coordinates
y = np.linspace(-2, 2, 17)     # Locations of y-coordinates
xx, yy = np.meshgrid(x, y)    # meshgrid returns two 2D arrays
u = 1 - yy**2                 # dx/dt = 1 - y**2
v = yy - xx                   # dy/dt = y - x
plt.quiver(x, y, u, v, color='b') # Plot the slope field
plt.xlabel('x')               # Add a label for the x-axis
plt.ylabel('y')               # Add a label for the y-axis
plt.title('Quiver plot')     # Add a title
plt.show()                   # Finally, show the figure
```



7.9 Stream Plots

The function **streamplot** plots the streamlines of a vector flow. This is similar to a slope field, but with the arrows of the slope field connected into continuous curves.

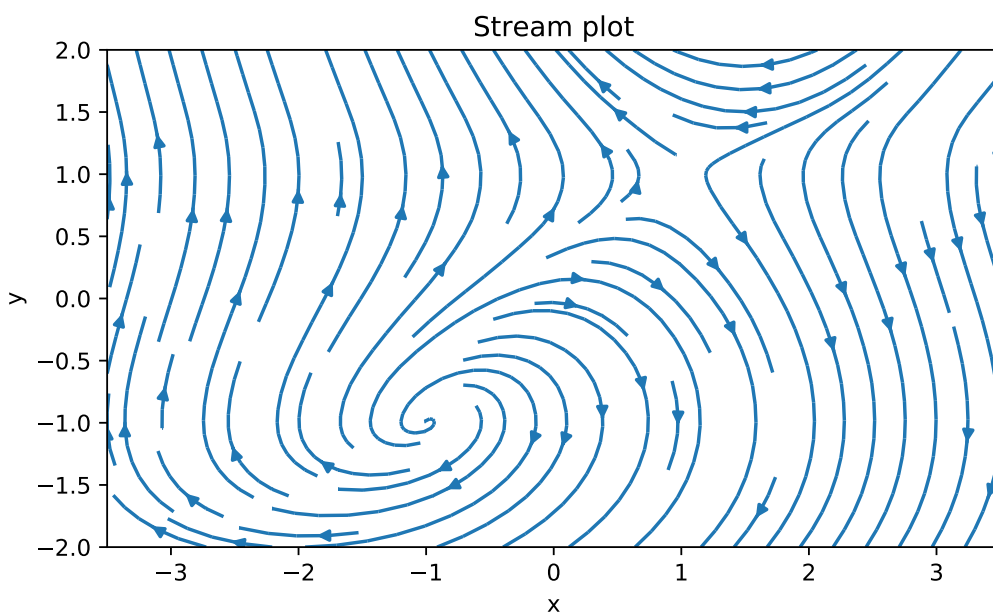
- The first two arguments are one dimensional arrays representing the x- and y-coordinates of the vectors.
- The next two arguments are two dimensional arrays representing the x- and y-components of the vectors.
- Keyword arguments allow the density, color, and thickness of the streamlines to be customized.

The following example is of a stream plot of the same system of ordinary differential equations used in the previous slope field example.

Plotting a Stream Plot

```
import numpy as np          # Imports linspace, meshgrid
import matplotlib.pyplot as plt # Import plotting functions

x = np.linspace(-3.5, 3.5, 15) # Locations of x-coordinates
y = np.linspace(-2, 2, 9)      # Locations of y-coordinates
xx, yy = np.meshgrid(x, y)    # meshgrid returns two 2D arrays
u = 1 - yy**2                 # dx/dt = 1 - y**2
v = yy - xx                   # dy/dt = y - x
plt.streamplot(x, y, u, v)    # Plot the streamplot
plt.xlabel('x')               # Add a label for the x-axis
plt.ylabel('y')               # Add a label for the y-axis
plt.title('Stream plot')     # Add a title
plt.show()                    # Finally, show the figure
```



7.10 Multiple Plots

The function **subplot** is used to plot multiple graphs on a single figure. This divides a figure into a grid of rows and columns, with plotting done in the currently active subplot.

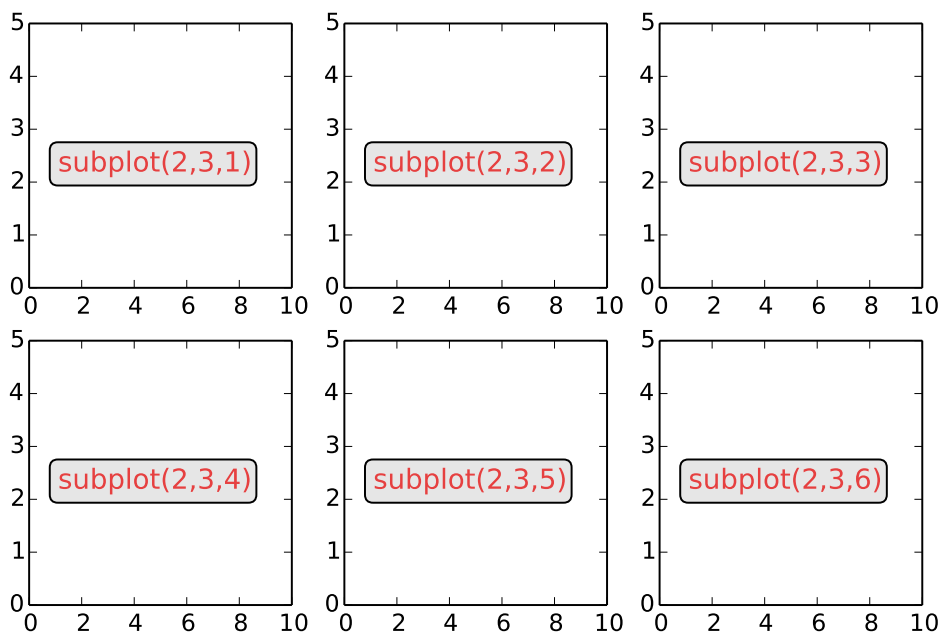
- Calls to subplot specify the number of rows, number of columns, and subplot number.
- Subplots are numbered from left to right, row by row, starting with 1 in the top left.
- All plotting is done in the location specified in the most recent call to subplot.
- If there are less than 10 rows, columns and subplots, subplot can be called with a string argument. For example, subplot(2, 3, 4) is the same as subplot("234").

The example below uses 2 rows and 3 columns. The "subplot" calls displayed on the figure show which call corresponds to each grid location.

Displaying Multiple Plots with **subplot**

```
import matplotlib.pyplot as plt      # Import plotting functions

fig=plt.figure(figsize=(8,5))       # Set the figure dimensions
nrows=2                             # Number of rows
ncols=3                             # Number of columns
for i in range(nrows*ncols):
    plt.subplot(nrows,ncols,i+1)    # Subplot numbering starts at 1
```



7.11 Formatting Text

The function `text` is used to add a text string to a plot at a given position.

- The first three positional arguments specify the x-position, y-position, and text string.
- The `fontsize` argument specifies the size of the font to use.
- The `fontstyle` argument specifies the style of font to use ('normal', 'italic' etc).
- The `fontweight` argument specifies how heavy the font should be ('normal', 'bold').
- The `family` argument specifies the font family to use ('serif', 'sans-serif' etc).
- The `color` argument specifies the color of the text.

These options can be combined together (for example, to specify text that is bold, red, italic and 14-point). The example below illustrates the use of these options.

Formatting Text

```
import matplotlib.pyplot as plt      # Import plotting functions
sizes = [10,12,14,16,18]
for sizepos, size in enumerate(sizes):
    plt.text(0, sizepos, "Font size = {}".format(size), fontsize=size)

styles = ['normal', 'italic', 'oblique']
for stylepos, style in enumerate(styles):
    plt.text(1, stylepos, "Style = {}".format(style), fontstyle=style)

families = ['serif', 'sans-serif', 'monospace']
for familypos, family in enumerate(families):
    plt.text(2, familypos, "Family = {}".format(family), family=family)

weights = ['normal', 'bold']
for weightpos, weight in enumerate(weights):
    plt.text(3, weightpos, "Weight = {}".format(weight), fontweight=weight)

colors = ['r', 'g', 'b', 'y', 'c']
for colorpos, color in enumerate(colors):
    plt.text(4, colorpos, "Color = {}".format(color), color=color)
```

| | | | | |
|----------------|-----------------|---------------------|----------------------|-----------|
| Font size = 18 | | | | Color = c |
| Font size = 16 | | | | Color = y |
| Font size = 14 | Style = oblique | Family = monospace | | Color = b |
| Font size = 12 | Style = italic | Family = sans-serif | Weight = bold | Color = g |
| Font size = 10 | Style = normal | Family = serif | Weight = normal | Color = r |

7.12 Formatting Mathematical Expressions

\LaTeX provides a way to format mathematical expressions in Matplotlib graphs in a similar way to Jupyter Notebook Markdown cells.

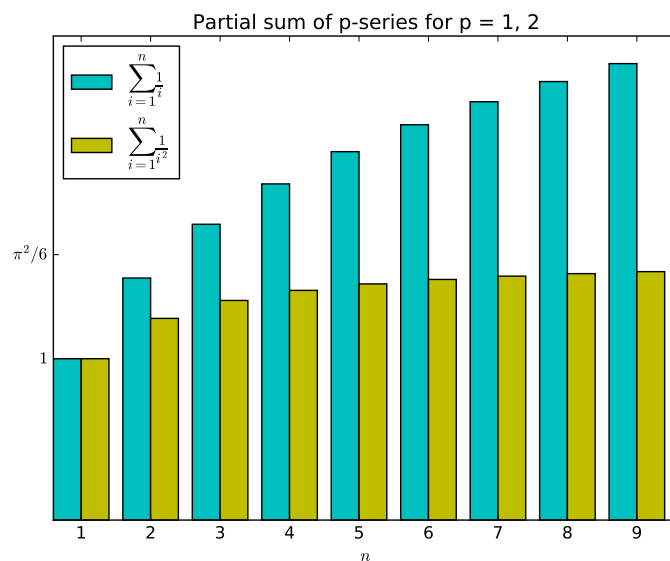
- Mathematical expressions are identified using `r"$\langle formula \rangle$"`.
- The syntax for $\langle formula \rangle$ is the same as that described in section 1.4.3 on \LaTeX .
- These expressions can be used anywhere a string is used, such as titles, axis and tick labels, and legends.

The example below illustrates several examples of mathematical expressions using \LaTeX .

Formatting Mathematical Expressions with \LaTeX

```
import numpy as np                # Import arange, cumsum, pi
import matplotlib.pyplot as plt   # Import plotting functions

x = np.arange(1,10)
y1 = np.cumsum(1/x)              # cumsum calculates the cumulative sum
y2 = np.cumsum(1/(x**2))
width = 0.4
plt.bar(x, y1, width=width, fc='c', label=r'\sum_{i=1}^n \frac{1}{i}')
plt.bar(x+width, y2, width=width, fc='y', label=r'\sum_{i=1}^n \frac{1}{i^2}')
ticks = x + width                # Shift the x-ticks to center on the bars
xlabels = [str(val) for val in x] # Labels must be strings
plt.xticks(ticks, xlabels)
ticks = [1, np.pi**2/6]
ylabels = [r'$1$', r'\pi^2/6$'] # Note that \pi renders as a symbol
plt.yticks(ticks, ylabels)
plt.xlabel(r'$n$')
plt.legend(loc='upper left')
plt.title('Partial sum of p-series for p = 1, 2')
```



8 Differential Equations

Differential equations of the form:

$$\frac{dy}{dt} = \text{func}(y, t, \dots)$$

can be solved in Python using **scipy.integrate.odeint**. The value of y can be a vector, which allows systems of equations and higher-order equations to all be solved by this one function.

- The first *func* argument is a function which computes the derivative of y at t .
- The second $y0$ argument specifies the initial state.
- The third t argument is an array of time points at which the solution is to be evaluated.
- The keyword *args* is a tuple of additional arguments to be passed to *func*.

The value returned by **odeint** is an array of solutions to the differential equation for each time in the array t . The first value in the solution is the initial value $y0$.

To apply **odeint** to an n -th order differential equation, we need to first solve for the n -th derivative:

$$\frac{d^n x}{dt^n} = f\left(t, x, \frac{dx}{dt}, \dots, \frac{d^{n-1}x}{dt^{n-1}}\right)$$

and then convert to a first-order system as follows:

$$\mathbf{y} = \left[x, \frac{dx}{dt}, \frac{d^2x}{dt^2}, \dots, \frac{d^{n-1}x}{dt^{n-1}} \right]$$

So:

$$\frac{d}{dt}\mathbf{y} = \left[\frac{dx}{dt}, \frac{d^2x}{dt^2}, \dots, \frac{d^n x}{dt^n} \right] = \left[\frac{dx}{dt}, \frac{d^2x}{dt^2}, \dots, f\left(t, x, \frac{dx}{dt}, \dots, \frac{d^{n-1}x}{dt^{n-1}}\right) \right]$$

The function defined by *func* then needs to accept the array \mathbf{y} as an argument and return an array of derivatives corresponding to the elements of \mathbf{y} .

8.1 First-Order Differential Equations

The following example solves the logistic growth equation:

$$\frac{dP}{dt} = kP(M - P)$$

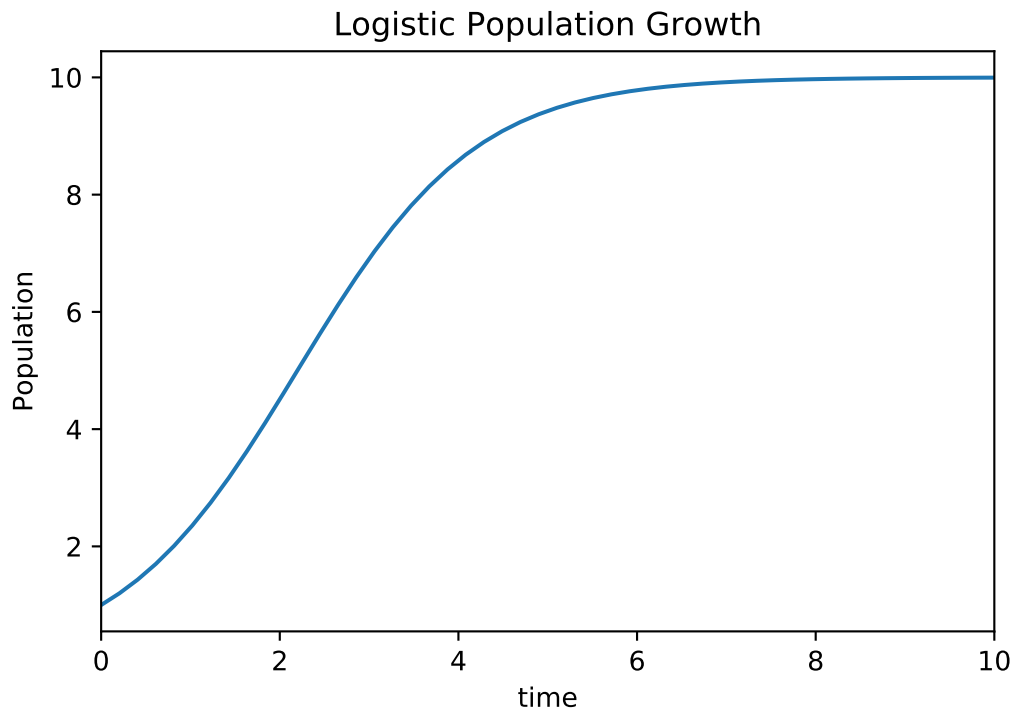
where P is the population at time t , k controls the rate of growth, and M is the maximum population size.

Solving a First-Order Differential Equation Using odeint

```
import numpy as np                    # Import linspace
import matplotlib.pyplot as plt      # Import plotting functions
from scipy.integrate import odeint

def logistic(P, t, k, M):            # Calculates derivative values
    dPdt = k*P*(M - P)
    return dPdt

P0 = 1                                # Initial value for P
t = np.linspace(0, 10)              # Array of time values
k = 0.1                              # Controls growth rate
M = 10                               # Maximum population size
P = odeint(logistic, P0, t, args=(k, M)) # Call odeint to solve the ODE
plt.plot(t, P)                       # Plot the solution against time
plt.xlabel('time')                   # Add a label for the x-axis
plt.ylabel('Population')             # Add a label for the y-axis
plt.title('Logistic Population Growth') # Add a title
plt.show()                           # Finally, show the figure
```



8.2 Higher Order Linear Equations

The following example solves the damped harmonic oscillator equation:

$$m \frac{d^2x}{dt^2} + c \frac{dx}{dt} + kx = 0$$

where x is the position of a body at time t , m is the mass of the body, c is the damping constant, and k is the spring constant. Because this is a second-order equation, to use **odeint** we define the vector $\mathbf{y} = [x, dx/dt] = [x, v]$. The derivative of \mathbf{y} is then given by:

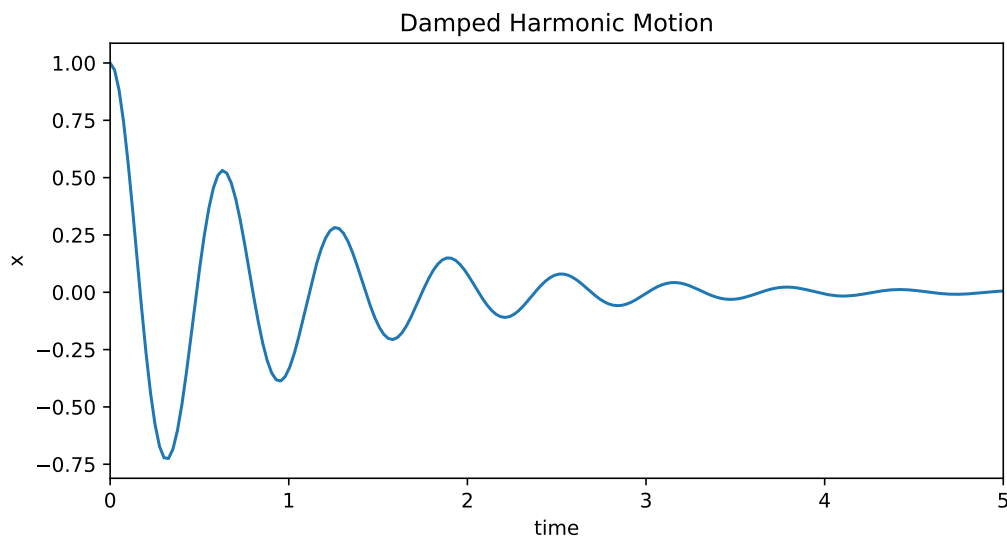
$$\frac{d\mathbf{y}}{dt} = \left[\frac{dx}{dt}, \frac{dv}{dt} \right] = \left[v, -\frac{cv + kx}{m} \right]$$

Solving a Second-Order Differential Equation Using odeint

```
import numpy as np                    # Import linspace
import matplotlib.pyplot as plt      # Import plotting functions
from scipy.integrate import odeint

def damped(y, t, m, c, k):           # Calculates derivative values
    x, v = y
    a = -(c*v + k*x)/m
    dydt = np.array([v, a])
    return dydt

y0 = np.array([1, 0])                # Initial position and velocity
t = np.linspace(0, 5, 200)          # Array of time values
m, c, k = 0.5, 1, 50                # Mass, damping, spring constants
y = odeint(damped, y0, t, args=(m, c, k)) # Call odeint to solve the ODE
plt.plot(t, y[:, 0])                 # Plot the solution against time
plt.xlabel('time')                   # Add a label for the x-axis
plt.ylabel('x')                       # Add a label for the y-axis
plt.title('Damped Harmonic Motion')  # Add a title
plt.show()                            # Finally, show the figure
```



8.3 Systems of Equations

The following example solves a system of equations known as a predator-prey model. The state is defined by the population of predators y and their prey x , coupled by the pair of equations:

$$\frac{dx}{dt} = x(a - py) \quad \frac{dy}{dt} = y(-b + qx)$$

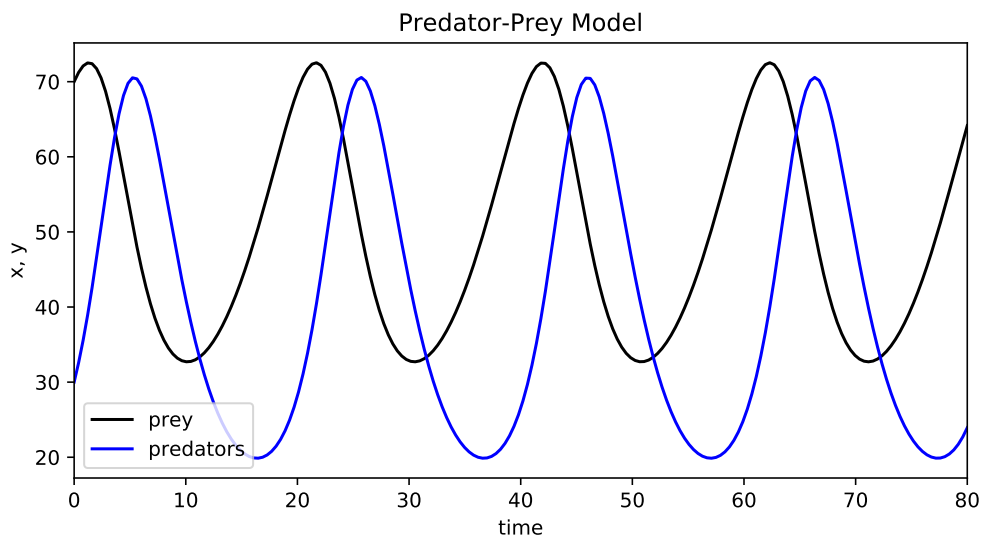
The constants a , p , b and q control rates of growth and decline due to interactions between the species, and to natural rates of birth and death.

Solving a System of Equations Using odeint

```
import numpy as np                                # Import linspace
import matplotlib.pyplot as plt                  # Import plotting functions
from scipy.integrate import odeint

def predator_prey(pp, t, a, p, b, q):           # Calculates derivative values
    x, y = pp
    dxdt = x*(a - p*y)
    dydt = y*(-b + q*x)
    dpp_dt = np.array([dxdt, dydt])
    return dpp_dt

pp0 = np.array([70, 30])                        # Initial populations
t = np.linspace(0, 80, 200)                    # Array of time values
args = (0.2, 0.005, 0.5, 0.01)                # Growth and decline rates
pp = odeint(predator_prey, pp0, t, args=args)   # Call odeint to solve the system
plt.plot(t, pp[:, 0], 'k', label='prey')       # Plot the prey against time
plt.plot(t, pp[:, 1], 'b', label='predators')  # Plot the predators against time
plt.xlabel('time')                             # Add a label for the x-axis
plt.ylabel('x, y')                             # Add a label for the y-axis
plt.title('Predator-Prey Model')               # Add a title
plt.legend()                                    # Add a legend
plt.show()                                     # Finally, show the figure
```



9 Additional Topics

9.1 Loading Numerical Files

We often need to load files containing numerical data into a NumPy array for further processing and display. Such data files typically consist of:

- Header information. This describes what the data represents and how it is formatted.
- A set of rows of numerical data. Each row contains the same number of values, separated by some string such as a comma or tab.

The NumPy function `numpy.loadtxt` can be used to load such data. This returns a NumPy array, where each row corresponds to a line in the data file. The first argument to this function is the data file name. Some of the optional keyword arguments are shown below.

- *dtype*. This is the data type of values in the array, which are floats by default.
- *delimiter*. This is the string used to separate values in each row. By default, any whitespace such as spaces or tabs are considered delimiters.
- *skiprows*. This is the number of rows to ignore at the start of the file before reading in data. It is usually used to skip over the header information, and defaults to 0.

The example shown below uses a file called “weather.dat”, which contains the following:

| Day | High-Temp | Low-Temp |
|-----|-----------|----------|
| 1 | 77 | 56 |
| 2 | 79 | 62 |

```
from numpy import loadtxt
```

*Import the **loadtxt** function.*

```
data = loadtxt("weather.dat", dtype=int,  
              skiprows=1)  
print(data)
```

```
[[ 1 77 56]  
 [ 2 79 62]]
```

*Load the “weather.dat” file, skipping the first row, and creating a 2×3 array of integers. For floats, the *dtype* argument would not be used.*

9.2 Images

Matplotlib provides functions for saving, reading, and displaying images. These images are either 2- or 3-dimensional NumPy arrays. In both cases, the first two axes of the array correspond to the rows and columns of the image. The third axis corresponds to the color of the pixel at each (column, row) coordinate.

- For a 2D array, the array values are floats in the range [0, 1]. These represent the luminance (brightness) of a grayscale image from black (0) to white (1).
- For a 3D array, the third axis can have either 3 or 4 elements. In both cases, the first three elements correspond to the red, green, and blue components of the pixel color. These can be either floats in the range [0, 1], or 8-bit integers of type 'uint8'. A fourth element corresponds to an “alpha” value representing transparency.

The main functions we use are:

`imread` Read an image file into an array.

`imsave` Save an image to file.

`imshow` Display an image array.

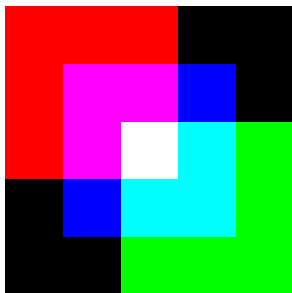
The following example creates an image as a 3D NumPy array of floats. The red, green and blue color components of the image are then set directly using array slicing.

Image Files: Creating a NumPy Image Array

```
import numpy as np          # Import zeros
import matplotlib.pyplot as plt # Import imsave, imread, imshow

img = np.zeros((100, 100, 3)) # Create an image array of 100 rows and columns.
img[:60,:60,0] = 1.          # Set the top-left corner to red.
img[40:,40:,1] = 1.         # Set the lower-right corner to green.
img[20:80,20:80,2] = 1.     # Set the center square to blue.

plt.imsave("squares.png", img) # Save the img array to the "squares.png" file
img2 = plt.imread("squares.png") # Read the file back to the img2 array
plt.imshow(img2)              # Display the image
```



9.3 Animation

An animation consists of a sequence of frames which are displayed one after the other. Animation using Matplotlib essentially involves updating the data associated with some drawn object or objects (such as points or lines), and redrawing these objects. Producing an animation therefore involves the following steps:

- Set up the variables and data structures relating to the animation.
- Draw the first frame.
- Repeatedly update the frame with new data.

Animations are generated using `FuncAnimation` from the `matplotlib.animation` module. This takes the following required arguments:

- *fig*. This is the figure in which the animation is to be drawn. It can be obtained using either the Matplotlib `figure` or `subplots` functions.
- *func*. This specifies the function to call to perform a single step of the animation. It should take a single argument which is the frame number (an integer). The frame number is used to update the values of drawn objects such as points or lines. If the *blit* keyword argument is `True`, this function should return a tuple of the modified objects that need to be redrawn.

`FuncAnimation` also takes the following keyword arguments.

- *frames*. An integer specifying the number of frames to generate.
- *init_func*. This specifies the function which is called once at the start to draw the background that is common to all frames. If the *blit* keyword argument is `True`, this function should also return a tuple of the modified objects that need to be redrawn.
- *interval*. This argument specifies the time (in ms) to wait between drawing successive frames.
- *blit*. If `True`, the animation only redraws the parts of the plot which have changed. This can help ensure that successive frames are displayed quickly.
- *repeat*. If `True` (the default), the animation will repeat from the beginning once it is finished.

The following example for Jupyter Notebook animates a point circling the origin with constant angular velocity. The `animate` function is defined to update the position of the point based on the frame number.

Animation: A Point Circling the Origin

```

%matplotlib                                     # Not %matplotlib inline
import numpy as np
import matplotlib.pyplot as plt
from matplotlib import animation

omega = .02                                     # Angular velocity
fig, ax = plt.subplots(figsize=(4,4))          # Get the figure & axes for the plot
ax.set_aspect('equal')                          # Make the axes have the same scale
point, = plt.plot([], [], 'ro', ms=10)        # "point" is the object drawn by plot
plt.xlim(-1.5,1.5)                             # - note that "plot" returns a tuple
plt.ylim(-1.5,1.5)                             # Set limits for the entire animation

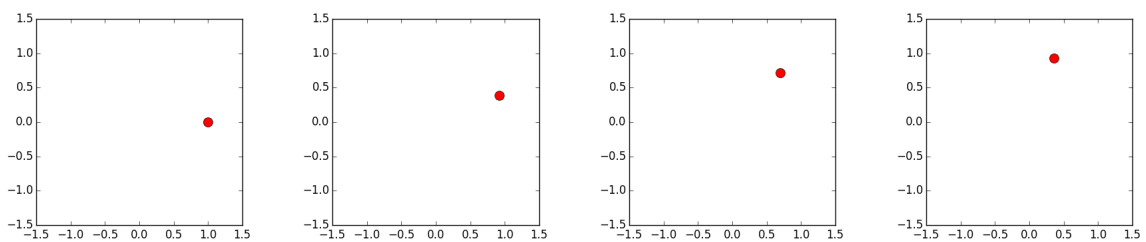
# Initialization function. This is called once to plot the background.
def init():
    point.set_data([], [])
    return point,                               # Return a tuple of the modified objects

# Animation function. This is called once per animation step.
# The integer i is the frame number.
def animate(i):
    x = np.cos(i*omega)
    y = np.sin(i*omega)
    point.set_data(x, y)                       # Update the x, y coordinates of the point
    return point,                               # Return a tuple of the modified objects

# Start the animator with a call to "FuncAnimation"
animation.FuncAnimation(fig, animate, init_func=init, frames=100, interval=20)

```

Some frames from this animation are shown below.



Note that in Jupyter Notebook the IPython magic we need to use is `%matplotlib` rather than `%matplotlib inline`. Inline graphs in Jupyter Notebook are static, meaning that once drawn, they cannot be updated. Using `%matplotlib` generates graphs in a separate window, where the updated data can be displayed.

9.4 Random Number Generation

NumPy provides a library of functions for random number generation in the `numpy.random` module. These return either a sample, or an array of samples of a given size, drawn from a given probability distribution. The main functions we use are:

- `random.rand` Samples are drawn from a uniform distribution over $[0, 1)$.
- `random.randint` Samples are integers drawn from a given range.
- `random.randn` Samples are drawn from the “standard normal” distribution.
- `random.normal` Samples are drawn from a normal (Gaussian) distribution.
- `random.choice` Samples are drawn from a given list or 1D array.

The following examples illustrate the use of these functions.

| | |
|--|--|
| <pre>import numpy as np print(np.random.rand())</pre> <hr/> <pre>0.723812203628</pre> | <p>Use random.rand to generate a single number uniformly drawn from the interval $[0, 1)$.</p> |
| <pre>print(np.random.rand(3))</pre> <hr/> <pre>[0.74654564 0.58764797 0.15557362]</pre> | <p>Use random.rand to generate an array of 3 random numbers drawn from $[0, 1)$.</p> |
| <pre>print(np.random.rand(2, 3))</pre> <hr/> <pre>[[0.65382707 0.71701863 0.5738609] [0.22064692 0.57487732 0.5710538]]</pre> | <p>Use random.rand to generate a 2×3 array of random numbers drawn from $[0, 1)$.</p> |
| <pre>print(np.random.randint(7))</pre> <hr/> <pre>3</pre> | <p>Use random.randint to generate an integer drawn from $\{0, \dots, 6\}$.</p> |
| <pre>print(np.random.randint(5,9,size=(2,4)))</pre> <hr/> <pre>[[5 5 5 8] [7 8 7 6]]</pre> | <p>Use random.randint to generate a 2×4 array of integers drawn from $\{5, 6, 7, 8\}$.</p> |
| <pre>print(np.random.randn(3))</pre> <hr/> <pre>[0.47481788 -0.7690172 0.42338774]</pre> | <p>Use random.randn to generate an array of samples drawn from the “standard” normal distribution.</p> |
| <pre>print(np.random.normal(100, 15))</pre> <hr/> <pre>111.676554337</pre> | <p>Use random.normal to generate a sample drawn from a normal distribution with $\mu = 100$, $\sigma = 15$.</p> |

9.5 Sound Files

Sound is a vibration that propagates through a medium such as air as a wave of pressure and displacement. Recording devices such as microphones convert this wave to an electrical signal. This signal is then sampled at regular intervals and converted to a sequence of numbers, which correspond to the wave amplitude at given times.

The WAV file format is a standard for storing such audio data without compression. WAV files contain two main pieces of information:

- The rate at which the wave has been sampled, usually 44,100 times per second.
- The audio data, usually with 16 bits used per sample. This allows $2^{16} = 65,536$ different amplitude levels to be represented.

The module `scipy.io.wavfile` provides functions to read and write such files.

`scipy.io.wavfile.read` Read a WAV file, returning the sample rate and the data.

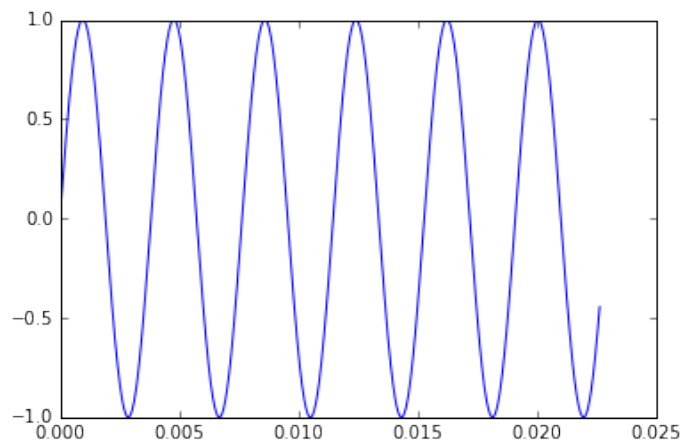
`scipy.io.wavfile.write` Write a NumPy array as a WAV file.

The following example creates and saves a WAV file with a single frequency at middle C, then plots the first 1000 samples of the data.

WAV File: Middle C

```
import numpy as np                    # Import linspace, sin
import matplotlib.pyplot as plt      # Import plotting functions
from scipy.io import wavfile

rate = 44100                          # Number of samples/second
end = 10                              # The file is 10 seconds long
time = np.linspace(0, end, rate*end+1) # Time intervals are 1/rate
freq = 261.625565                    # Frequency of "middle C"
data = np.sin(2*np.pi*freq*time)     # Generate the sine wave
wavfile.write("middleC.wav", rate, data) # Write the array to a WAV file
plt.plot(time[:1000], data[:1000])   # Plot the first 1000 samples
plt.show()                            # Finally, show the figure
```



9.6 Linear Programming

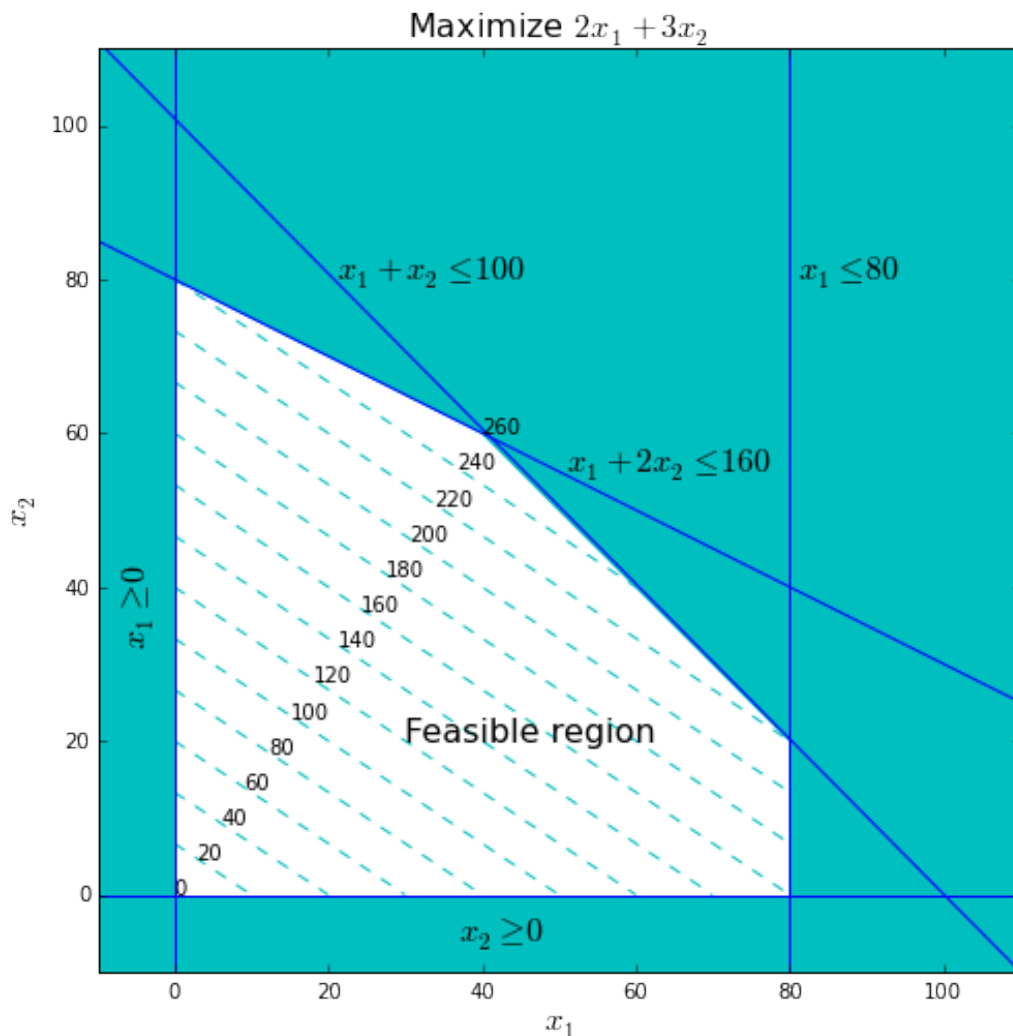
Linear programming problems are a special class of optimization problem. They involve finding the maximum (or minimum) of some linear objective function $f(\mathbf{x})$ of a vector of variables $\mathbf{x} = (x_1, x_2, \dots, x_n)$, subject to a set of linear equality and inequality constraints.

Since the objective function and constraints are linear, we can represent the problem as:

Maximize $\mathbf{c}^T \mathbf{x}$, where the vector \mathbf{c} contains the coefficients of the objective function, subject to $\mathbf{A}_{ub} * \mathbf{x} \leq \mathbf{b}_{ub}$, where \mathbf{A}_{ub} is a matrix and \mathbf{b}_{ub} a vector, and $\mathbf{A}_{eq} * \mathbf{x} = \mathbf{b}_{eq}$, where \mathbf{A}_{eq} is a matrix and \mathbf{b}_{eq} a vector.

An example of such a problem would be: $\mathbf{x} = \{x_1, x_2\}$. Maximize $f(\mathbf{x}) = 2x_1 + 3x_2$ subject to the inequality constraints (i) $0 \leq x_1 \leq 80$, (ii) $x_2 \geq 0$, (iii) $x_1 + x_2 \leq 100$, and (iv) $x_1 + 2x_2 \leq 160$.

This example is graphed below, showing the level curves of $f(\mathbf{x})$.



The function `scipy.optimize.linprog` implements the "simplex algorithm" we discuss in class to solve this problem. The arguments to this function are the values \mathbf{c} , \mathbf{A}_{ub} , \mathbf{b}_{ub} , \mathbf{A}_{eq} and

\mathbf{b}_{eq} given above. An optional *bounds* argument represents the range of permissible values that the variables can take, with `None` used to indicate no limit.

Applying **linprog** to this problem is done as shown below.

Linear Programming: Finding the Maximum Value

```
import numpy as np
from scipy.optimize import linprog

c = np.array([-2, -3])           # Negative coefficients of f(x)
A_ub = np.array([[1, 1], [1, 2]]) # Matrix of the inequality coefficients
b_ub = np.array([100, 160])     # Vector of the inequality upper bounds
bounds = [(0, 80), (0, None)]   # Each tuple is a (lower, upper) bound
result = linprog(c, A_ub=A_ub, b_ub=b_ub, bounds=bounds)
print(result.x)                 # The result "x" field holds the solution
```

This yields the correct solution for x_1 and x_2 , as seen in the graph above:

```
[ 40.  60.]
```



Note that **linprog** finds the minimum of $f(\mathbf{x})$. To find the maximum, the negative of the **c** coefficient values needs to be used instead.

IV Programming Tips

10 Programming Style

This chapter contains some tips on how to make programs easier to read and understand. Programs are written first and foremost to be understood by human beings, not by computers. Ideally, it should be possible a year from now for you to pick up the code that you're writing today and still understand what you were doing and why.

10.1 Choosing Good Variable Names

Good variable names make reading and debugging a program much easier. Well chosen names are easy to decipher, and make the intent clear without additional comments.

- A variable name should fully and accurately describe the data it represents. As an example, `date` may be ambiguous whereas `current_date` is not. A good technique is to state in words what the variable represents, and use that for the name.
- Names that are too short don't convey enough meaning. For example, using `d` for date or `cd` for current date is almost meaningless. Research shows that programs with variable names that are about 9 to 15 characters long are easiest to understand and debug.
- Variable names should be problem-oriented, referring to the problem domain, not how the problem is being solved. For example, `planet_velocity` refers to the problem, but `vector_3d` refers to how this information is being represented.
- Loop indices are often given short, simple names such as `i`, `j` and `k`. This is okay here, since these variables are just used in the loop, then thrown away.
- If loops are nested, longer index names such as `row` and `column` can help avoid confusion.
- Boolean variables should have names that imply either `True` or `False`. For example, `prime_found` implies that either a prime has been found, or it hasn't.
- Boolean variables should be positive. For example, use `prime_found` rather than `prime_not_found`, since negative names are difficult to read (particularly if they are negated).
- Named constants should be in uppercase and refer to what the constant represents rather than the value it has. For example, if you want to use the same color blue for the font in every title, then define the color in one place as `TITLE_FONT_COLOR` rather than `FONT_BLUE`. If you later decide to have red rather than blue titles, just redefine `TITLE_FONT_COLOR` and it keeps the same meaning.

10.2 Choosing Good Function Names

The [recommended style](#) for naming functions in Python is to use all lowercase letters, separated by underscores as necessary. As with variable names, good function names can help make the intent of the code much easier to decipher.

- For procedures (functions that do something and don't return a value), use a verb followed by an object. An example would be `plot_prime_distribution`.
- For functions that return values, use a description of what the returned value represents. An example would be `miles_to_kilometers`.
- Don't use generic names such as `calculate_stuff` or numbered functions such as `function1`. These don't tell you what the function does, and make the code difficult to follow.
- Describe everything that the function does, and make the function name as long as is necessary to do so. If the function name is too long, it may be a sign that the function itself is trying to do too much. In this case, the solution is to use shorter functions which perform just one task.

10.3 No "Magic Numbers"

Magic numbers are numbers such as 168 or 9.81 that appear in a program without explanation. The problem with such numbers is that the meaning is unclear from just reading the number itself.

- Numbers should be replaced with named constants which are defined in one place, close to the start of your code file.
- Named constants make code more readable. It's a lot easier to understand what `HOURS_PER_WEEK` is referring to than the number 168.
- If a number needs to change, named constants allow this change to be done in one place easily and reliably.

10.4 Comments

It's not necessary to comment every line of code, and "obvious" comments which just repeat what the code does should be avoided. For example, the endline comment in the following code is redundant and does nothing to explain what the code is for.

```
x += 1 # Add 1 to x
```

Good comments serve two main purposes:

- "Intent" comments explain the purpose of the code. They operate at the level of the problem (*why* the code was written) - rather than at the programming-language level (*how* the code operates). Intent is often one of the hardest things to understand when reading code written by another programmer.

- "Summary" comments distill several lines of code into one or two sentences. These can be scanned faster than the code itself to quickly understand what the code is doing. For example, suppose you are creating several different graphs for a report. A summary comment before each set of plotting commands can describe which figure in the report the code is producing.

Endline comments are those at the end of a line, after the code. They are best avoided for a number of reasons.

- Endline comments are short by necessity as they need to fit into the space remaining on a line. This means that they tend to be cryptic and uninformative.
- Endline comments are difficult to keep aligned (particularly as the code changes), and if they're not aligned they become messy and interfere with the visual structure of the code.

A final note is to get in the habit of documenting code files. At the top of every file, include a block comment describing the contents of the file, the author, and the date the file was created. An example would be:

Sample File Header

```
# MTH 337: Intro to Scientific and Mathematical Computing, Spring 2017  
# Report 1: Primitive Pythagorean Triples  
# Created by Adam Cunningham 1/30/2017
```

10.5 Errors and Debugging

The following suggestions may help to reduce errors.

- Test each function completely as you go.
- In the initial stages of learning Python, test each few lines of code before moving on to the next.
- Add "print" statements inside a function to print out the intermediate values of a calculation. This can be used to check that a function is working as required, and can always be commented out afterwards.

In the event of an error being generated, IPython will typically give as much information as possible about the error. If this information is not sufficient, the `%debug` magic will start the IPython debugger. This lets the current values of variables inside a function be examined, and allows code to be stepped through one line at a time.

II Further Reading

The following books may prove useful for further study or reference.

- L. Felipe Martins. *IPython Notebook Essentials*. Packt Publishing Ltd, Birmingham. 2014.
A fairly short introduction to using NumPy and Matplotlib in Jupyter Notebooks. This is not a Python tutorial, although there is a brief review of Python in the appendix.
- Steve McConnell. *Code Complete: A Practical Handbook of Software Construction, Second Edition*. Microsoft Press. 2004.
A general guide to code writing and software construction, this book focuses on questions of software design rather than any specific language. More useful to an intermediate-level programmer who wants to improve their skills. No references to Python.
- Bruce E. Shapiro. *Scientific Computation: Python Hacking for Math Junkies*. Sherwood Forest Books, Los Angeles. 2015.
A tutorial for Python, NumPy and Matplotlib that also covers many of the same scientific and mathematical topics as this class.
- John M. Stewart. *Python for Scientists*. Cambridge University Press, Cambridge. 2014.
A good introduction to Python, NumPy, Matplotlib and three-dimensional graphics. Extensive treatment of numerical solutions to ordinary, stochastic, and partial differential equations.

Index

%%timeit magic, 12
L^AT_EX, 14, 77

abs, 17
and, 64
animate, 84
animation, 84
append, 28
arange, 55, 58, 62
around, 57
array creation, 54
astype, 58

bar, 68
bar plots, 68
Boolean expressions, 34
Boolean type, 18
break, 40

capitalize, 20
close, 42, 43
comments, 50, 92
complex numbers, 16
conditional expressions, 36
container, 26
continue, 40, 41
contour, 72
contour plots, 72
contourf, 72
copy, 55
count, 21, 28

def, 45
dictionaries, 32
dictionary comprehensions, 49
difference, 31
dir, 48
dtype, 53
dtype, 57

elif, 35, 36
else, 35
empty, 55
empty_like, 55
enumerate, 38
exp, 60
extend, 28

filled contour plots, 72
floats, 16
for, 37, 38, 40, 43, 50
format, 21
function names, 92
functions, 45
generator expressions, 50

hist, 70
histograms, 70

if, 34–36
if-else, 35
imag, 17
import, 48
index, 21, 28
insert, 28
integers, 16
intersection, 31
IPython magics, 12
items, 39

Jupyter Notebook, 9

lambda, 47
len, 26
line styles, 65
linear programming, 88
linprog, 89
list comprehensions, 49

- list methods, 28
- loadtxt**, 82
- logical_and**, 64
- logical_not**, 64
- logical_or**, 64
- lower**, 20

- magic numbers, 92
- Markdown, 13
- marker styles, 65
- math**, 49
- Matplotlib, 65
- max**, 29, 60
- meshgrid**, 56, 57
- min**, 29, 60
- modules, 48

- not**, 64
- NumPy, 53

- odeint**, 78, 80
- ones**, 55
- ones_like**, 55
- open**, 42, 43
- or**, 64

- pie**, 71
- pie chart, 71
- plot**, 69
- polar**, 69
- polar plots, 69
- pop**, 28
- print**, 19
- programming style, 91

- quiver**, 73

- random numbers, 86
- random.normal**, 86
- random.rand**, 86
- random.randint**, 60, 86
- random.randn**, 86
- range**, 38
- read**, 42, 43

- readlines**, 42, 43
- real**, 17
- remove**, 28
- replace**, 21
- reshape**, 58, 61
- return**, 46

- scipy.integrate.odeint**, 78
- scipy.io.wavfile**, 87
- sets, 30
- shape, 53
- show**, 66, 67
- sin**, 59
- slope field, 73
- sorted**, 48
- sound, 87
- split**, 21
- streamplot**, 74
- string indexing, 19
- string methods, 20
- string slicing, 19
- strings, 18
- subplot**, 75
- sum**, 60
- symmetric_difference**, 31

- text**, 76
- try-except**, 41, 42
- tuples, 29
- type casting, 22

- union**, 31
- upper**, 20

- values**, 39
- variable names, 24, 91

- WAV files, 87
- while**, 40

- zeros**, 55
- zeros_like**, 55
- zip**, 39