NUMERICAL METHODS WEEK 6

RECAP

Learning outcomes:

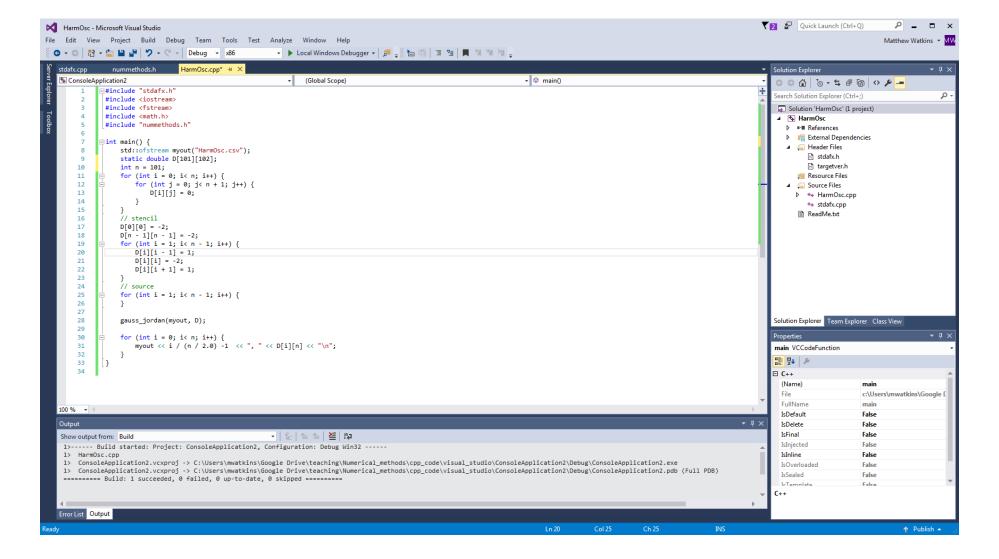
- Practice using some of the methods we have covered so far.
- See how to easily and tidyly reuse code.

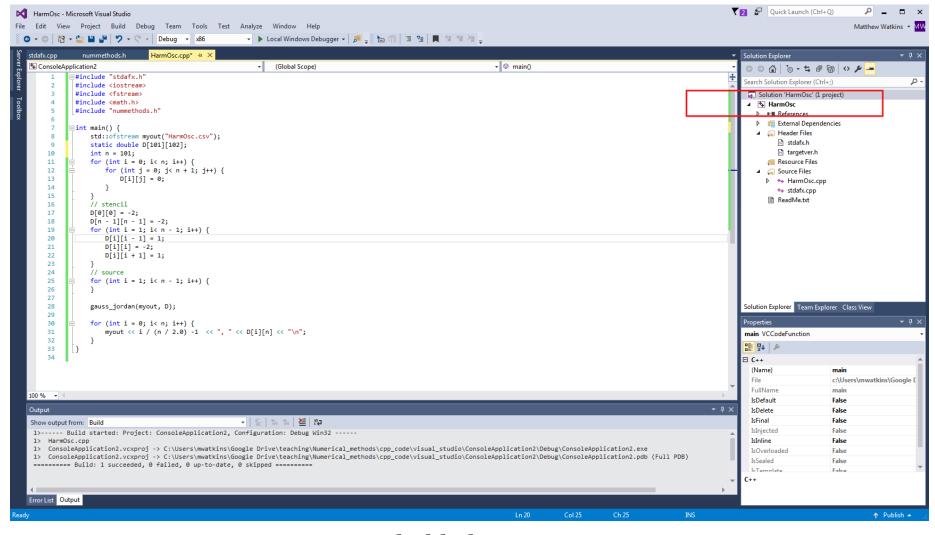
MATT WATKINS MWATKINS@LINCOLN.AC.UK

REUSING CODEThe property of the property o

The solution is to **include** other files or, generally, libraries.

We can make our own library file by copying functions other than main into a file called 'my_library.h'

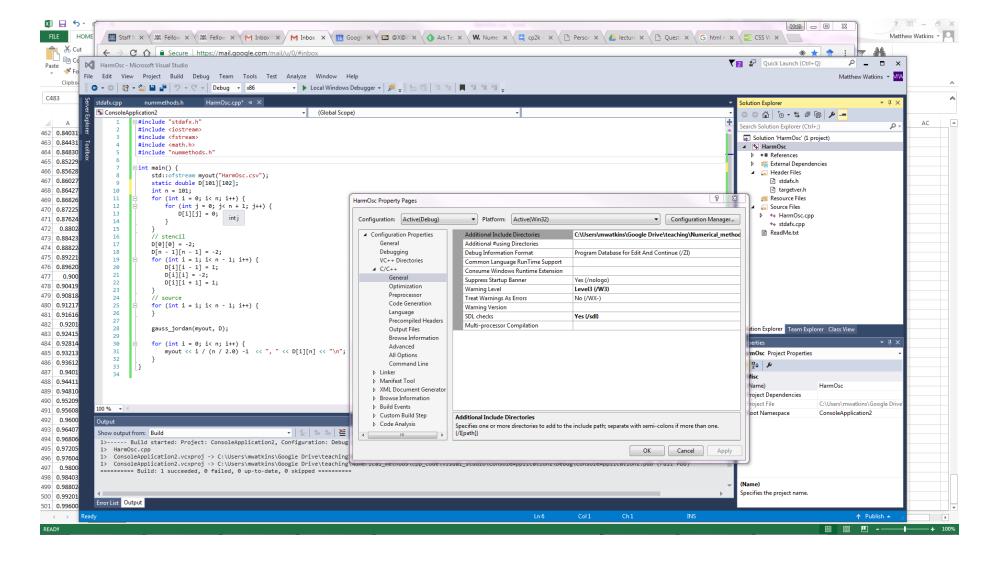




This is my example project. Right click on the **bolded** project name in the pane to right.

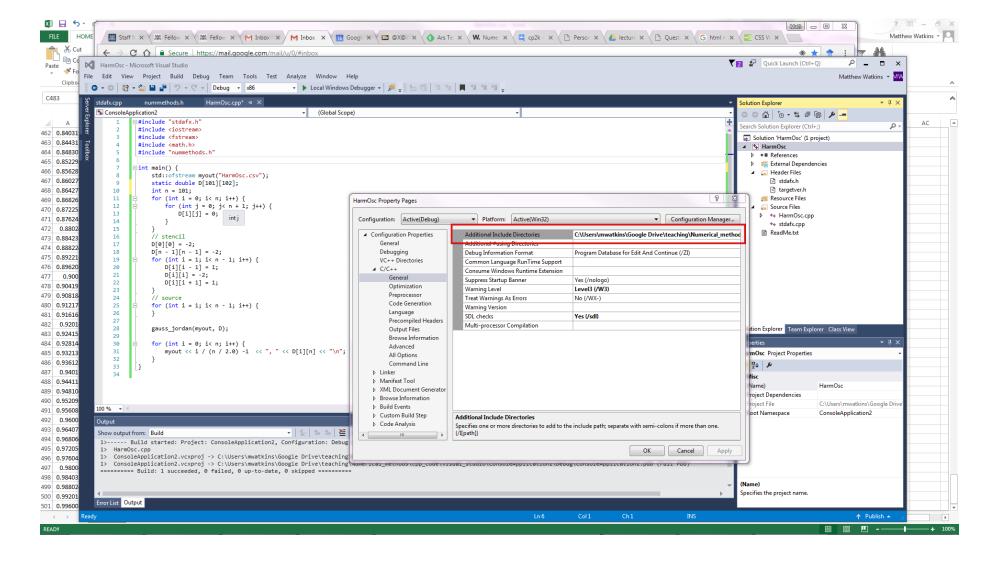
Then select properties.

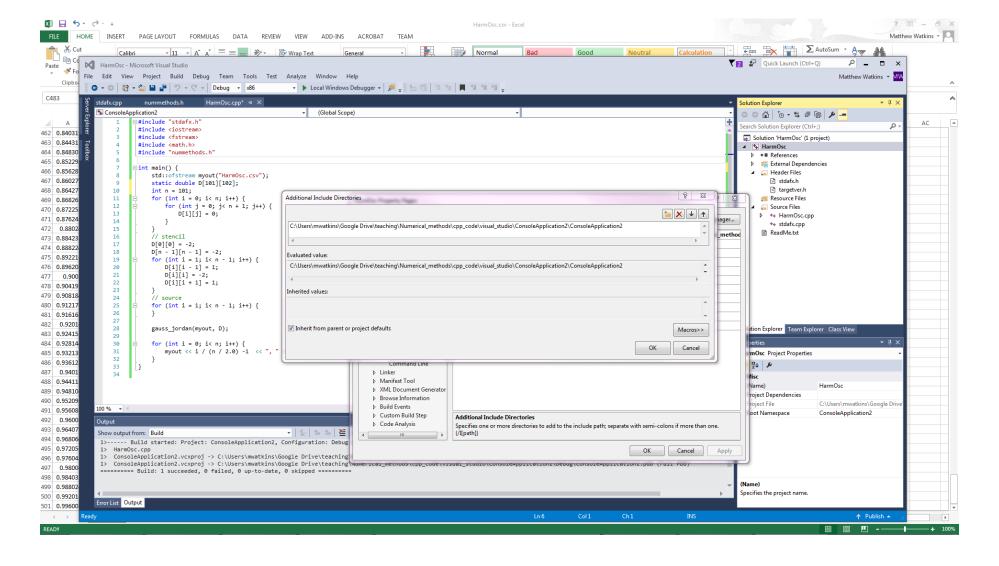
A new window should have appeared, like this:



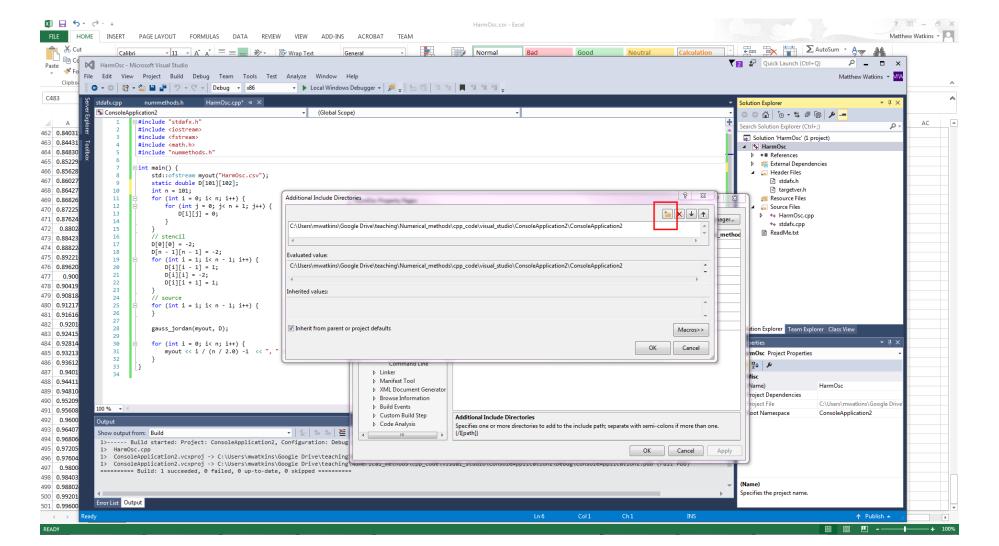
Open the C/C++ then the General tab

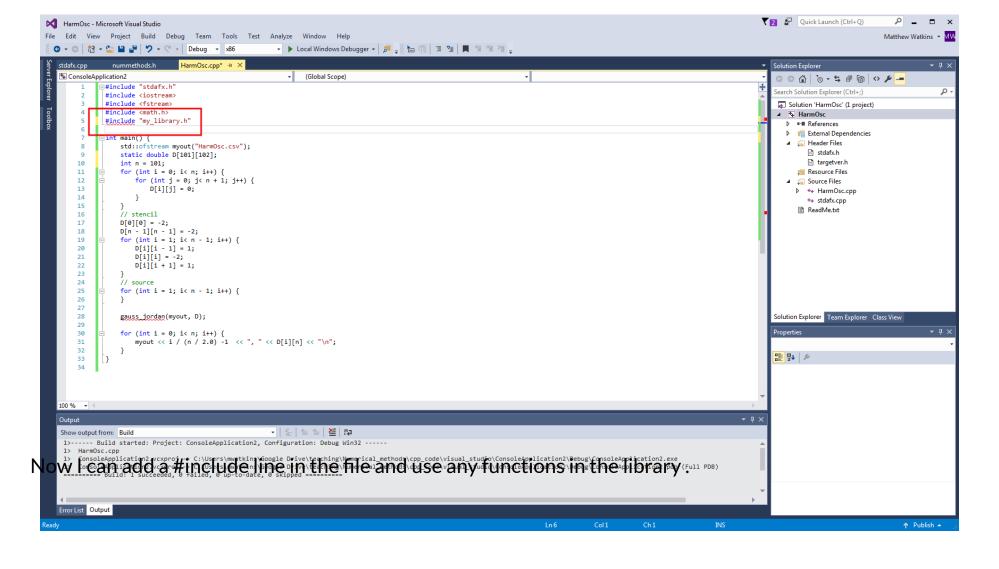
Then click on drop down button in the second column of the the 'Additional Include Directories' row and select edit





Click on the folder icon, and select the directory where you saved your 'my_library.h' file in





POWER METHOD

This is a way of finding the eigenvector with the largest eigenvector

- Start with a guess \mathbf{x}_0 at the eigenvector (it must not be orthogonal to the true eigenvector).
- ullet Update the eigenvector $\mathbf{x}_i = rac{\mathbf{A}\mathbf{x}_{n-1}}{||\mathbf{A}\mathbf{x}_{n-1}||}$

Find the largest eigenvector of

$$\mathbf{A} = egin{pmatrix} 1 & -1 & -1 & -1 \ -1 & 2 & 0 & 0 \ -1 & 0 & 3 & 1 \ -1 & 0 & 1 & 4 \end{pmatrix}.$$

using the power method.

We basically used this when studying Markov Chains in year 1.

GENERAL LINEAR LEAST SQUARES

Simple linear, polynomial and multiple linear regression can be generalised to the following linear leastsquares model $y_i(a_0,a_1,a_2,a_3\ldots a_n;x_0,x_1,x_2,x_3\ldots x_n)=a_0z_0+a_1z_1+a_2z_2+\cdots+a_mz_m+e$ are now not just coordinates x_i but some *predefined* functions of those positions $z_0(x_0,x_1,x_2,x_3\ldots x_n),z_1(x_0,x_1,x_2,x_3\ldots x_n),\ldots,z_m(x_0,x_1,x_2,x_3\ldots x_n)$ are m+1 basis functions.

The linear refers to the parameters a_0, a_1, \ldots, a_m , the zs can be highly non-linear

For instance.

$$y(a_0,a_1,a_2;x_0) = a_0 + a_1\cos(\omega x_0) + a_2\sin(\omega x_0)$$

fits this model with $z_0=1, z_1=\cos(\omega x_0)$ and $z_2=\sin(\omega x_0)$. Where ω is a single independent variable and ω is a predefined constant.

I'll supress the functional dependence notation as it gets too much. Remember the zs are predefined - only the as are optimized.

GENERAL LINEAR LEAST SQUARES EXAMPLE

Let's redo the quadratic fit example using the general linear least squares formalism. The general expression

$$y_i(a_0,a_1,a_2,a_3\dots a_n;x_0,x_1,x_2,x_3\dots x_n)=a_0z_0(x_i)+a_1z_1(x_i)+a_2z_2(x_i)+\dots+a_mz_m(x_i)+a_1z_1(x_i)+a_2z_2(x_i)+\dots+a_nz_m(x_i)$$

becomes

$$y_i(a_0,a_1,a_2;x_0,x_1,x_2,x_3\dots x_n) = a_0z_0(x_i) + a_1z_1(x_i) + a_2z_2(x_i) + e$$

as we only have three parameters.

Comparing to the formula for a quadratic fit

$$y_i(a_0,a_1,a_2;x_0,x_1,x_2,x_3\dots x_n)=a_0+a_1x_i+a_2x_i^2+e$$

we see that $z_0(x_i)=1$, $z_1(x_i)=x_i$ and $z_2(x_i)=x_i^2$

We write the set of equations for the all the n+1 data points in matrix form

$$egin{bmatrix} y_0 \ y_1 \ dots \ y_n \end{bmatrix} = egin{bmatrix} z_0(x_0) & z_1(x_0) & \cdots & z_m(x_0) \ z_0(x_1) & z_1(x_1) & \cdots & z_m(x_1) \ dots & dots & \ddots & dots \ z_0(x_n) & z_1(x_n) & \cdots & z_m(x_n) \end{bmatrix} egin{bmatrix} a_0 \ a_1 \ dots \ a_m \end{bmatrix} + egin{bmatrix} e(x_0) \ e(x_1) \ dots \ e(x_n) \end{bmatrix}$$

we'll define

$$\{Y\} = egin{bmatrix} y(x_0) \ y(x_1) \ dots \ y(x_n) \end{bmatrix}, [Z] = egin{bmatrix} z_0(x_0) & z_1(x_0) & \cdots & z_m(x_0) \ z_0(x_1) & z_1(x_1) & \cdots & z_m(x_1) \ dots & dots & \ddots & dots \ z_0(x_n) & z_1(x_n) & \cdots & z_m(x_n) \end{bmatrix}, \{e\} = egin{bmatrix} e(x_0) \ e(x_1) \ dots \ e(x_n) \end{bmatrix}$$

which contain the values of the basis functions and the errors at each point, and

$$\{A\} = egin{bmatrix} a_0 \ a_1 \ dots \ a_m \ . \end{bmatrix}$$

which contains the m+1 parameters we optimize.

We can also express error in our model as a sum of the squares much like before:

$$S_r = \sum_{i=0}^n \left(y_i - \sum_{j=0}^m a_j z_j(x_i)
ight)^2$$

You can show by taking partial derivatives that S_r is minimised when

$$[[Z]^T[Z]]\{A\} = \{[Z^T]\{Y\}\}$$

More details can of the derivation can be found here (http://fourier.eng.hmc.edu/e176/lectures/NM/node35.html), though the notation is a little different.

POWER METHOD, WHY DOES IT WORK?

If our matrix is diagonalizable it has a set of orthogonal eigenvectors $\nu_1, \nu_2, \dots, \nu_n$ with eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ with $|\lambda_1| > |\lambda_j|$ for j > 1.

Any vector can be expanded as a linear combination the ν_i :

$$b_0 = c_1 v_1 + c_2 v_2 + \cdots + c_m v_m.$$

If we apply our matrix ${f A}$ to b_0 k times we get

$$egin{array}{lll} A^k b_0 & = & c_1 A^k v_1 + c_2 A^k v_2 + \cdots + c_m A^k v_m \ & = & c_1 \lambda_1^k v_1 + c_2 \lambda_2^k v_2 + \cdots + c_m \lambda_m^k v_m \ & = & c_1 \lambda_1^k \left(v_1 + rac{c_2}{c_1} \left(rac{\lambda_2}{\lambda_1}
ight)^k v_2 + \cdots + rac{c_m}{c_1} \left(rac{\lambda_m}{\lambda_1}
ight)^k v_m
ight). \end{array}$$

using the fact that for the eigenvectors $\mathbf{A}
u_i = \lambda_i
u_i$.

Because λ_1 is larger than any other λ , all the terms except the first in the bracket tend to zero as k gets larger.

SUMMARY AND FURTHER READING

Ensure that you can do the material tested today - make and multiply matrices, solve linear equations - we will be using these repeatedly in the coming weeks.

You should be reading additional material to provide a solid background to what we do in class

I suggest starting with Chapra and Canale, but there are many other text books in the library - find the one that works for you.

SNAKE

Use the arrow keys

start game

JSXG	raph	v1.4	.6 C	opyri	ght (C) se	e htt	ps://j	sxgr	aph.d	org				