

THE CHINESE UNIVERSITY OF HONG KONG
Department of Mathematics

Mathematical Modeling Project Team
mathmodel@math.cuhk.edu.hk

Modeling with R Shiny (Multiple Regression)

Last updated: October 1, 2026

1 Introduction to the Multi-Regression R Shiny tool 多元線性迴歸R Shiny 工具簡介

R is a powerful statistical programming language that allows us to perform data analysis, which is useful for mathematical modeling. Even better, we have set up a set of online tools using R Shiny so that you can perform various mathematical modelling tasks without any coding background.

R 是一種功能強大的統計程式語言，讓我們能夠進行數據分析，對數學建模十分有用。更理想的是，我們利用R Shiny 建立了一系列線上工具，讓使用者即使不具備程式設計背景，也能夠執行各種數學建模任務。

In particular, multiple regression is a common modelling technique used to investigate the relationship between a response variable and several explanatory variables.

具體而言，多元線性迴歸是一種常見的建模方法，用於研究一個因變量與多個自變量之間的關係。

The model has the form
模型形式如下：

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p + \epsilon.$$

Here is the guideline on how to use our Multi-Regression R Shiny tool.

以下是有關如何使用我們的多元線性迴歸R Shiny 工具的指引。

2 Using the Multi-Regression R Shiny tool 使用多元線性迴歸R Shiny 工具

To use the tool, please visit our main website:

如要使用此工具，可前往我們的主網站：

<https://cuhkmathmodel.com>

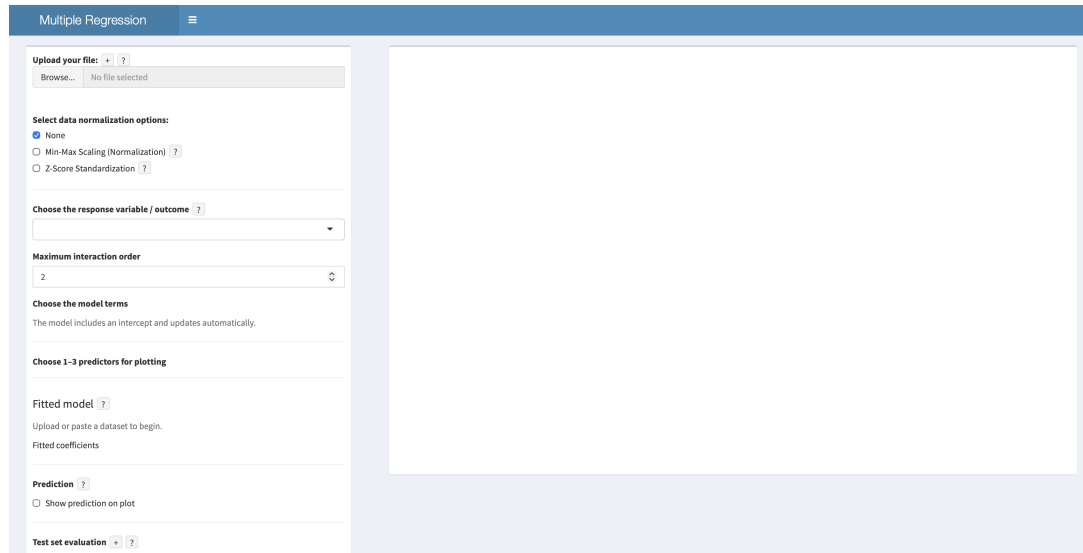
for the list of tools, or simply go to

瀏覽工具列表，或直接前往

<https://mathmodelcuhk.shinyapps.io/multi-regression/>

for the Multi-Regression R Shiny tool.
使用多元線性迴歸R Shiny 工具。

After getting into the website, you can see the following:
進入網站後，你會看到以下畫面：



The screenshot shows the 'Multiple Regression' Shiny app interface. On the left is a sidebar with various controls, and on the right is a large empty plot area.

Multiple Regression

Upload your file: [+](#) [?](#)
Browse... No file selected

Select data normalization options:
☒ None
☐ Min-Max Scaling (Normalization) [?](#)
☐ Z-Score Standardization [?](#)

Choose the response variable / outcome [?](#)
▼

Maximum interaction order
2

Choose the model terms
The model includes an intercept and updates automatically.

Choose 1-3 predictors for plotting

Fitted model [?](#)
Upload or paste a dataset to begin.
Fitted coefficients

Prediction [?](#)
☐ Show prediction on plot

Test set evaluation [+](#) [?](#)

Don't worry if you think this is complicated! We will guide you step by step, and you can do multiple regression using our R Shiny tool.
如果你覺得這個界面有點複雜，不用擔心！我們會逐步引導你，讓你能夠順利使用我們的R Shiny 工具進行多元線性迴歸。

2.1 Step 1: Data input

步驟一：數據輸入

At the top left-hand corner, you can see “Upload your file”, this is where we input the data. There are three different buttons.

在左上角，你會看到「Upload your file」，這就是我們輸入數據的地方。這裡有三個不同的按鈕。

Firstly, the “?” gives you the general guide on data input.

首先，「？」按鈕會提供關於數據輸入的一般指南。

Secondly, the “+” allows us to input data directly. After clicking it, you can see the following:

其次，「+」按鈕讓你可以直接輸入數據。點擊後，你會看到以下畫面：

Input your data

Paste comma-separated data with a unique column name in the first row for every variable.

```
"x","y","z"  
1.5,40.6,15.6  
3.8,65,25.8  
8.9,70.8,4.9  
10,78,3.56  
12,84,25  
20,108.5,29.3
```

CancelSubmit

On the first line, you can input the names of the variables respectively. For example, sales will be used as the response variable, while TV, radio and newspaper will be used as explanatory variables. Therefore, on the first line we type “TV,radio,newspaper,sales”

在第一行，你可以分別輸入變量名稱。例如，我們可以選擇**sales** 作為因變量。選擇**TV**、**radio** 及**newspaper** 作為自變量，以觀察它們之間的相關性。因此，在第一行我們輸入「TV,radio,newspaper,sales」。

For the following lines, you can simply type in the data points respectively. You should be able to have a result that looks like this:
在接下來的行數，你可以分別輸入各個數據點。你應該會得到類似這樣的結果：

Input your data

Paste comma-separated data with a unique column name in the first row for every variable.

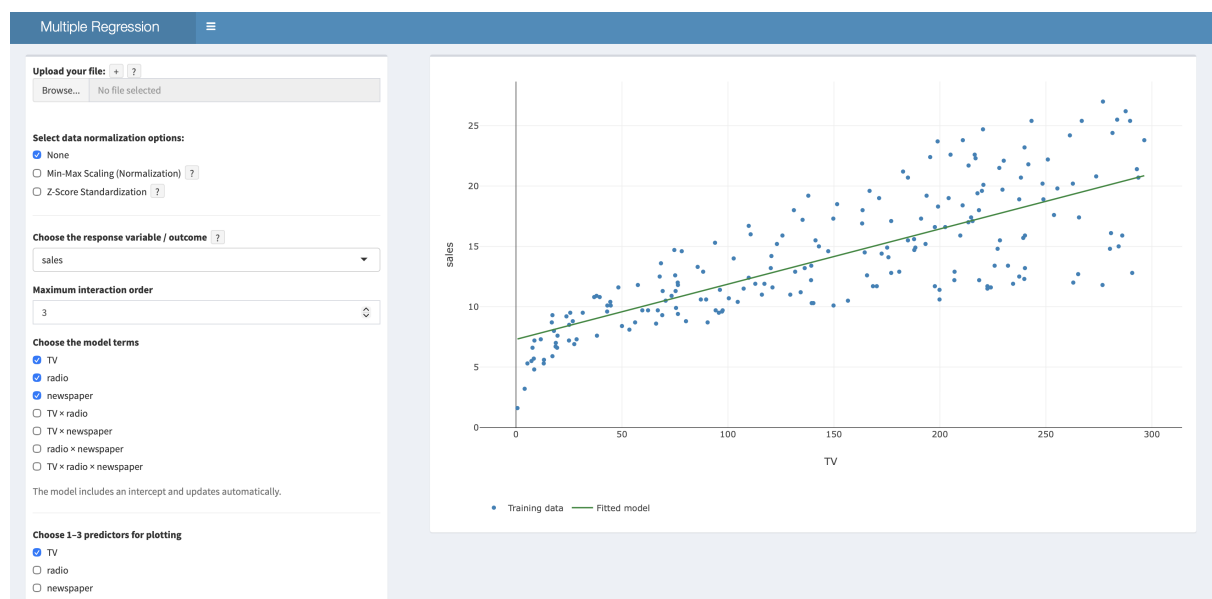
```
"TV","radio","newspaper","sales"  
230.1,37.8,69.2,22.1  
44.5,39.3,45.1,10.4  
17.2,45.9,69.3,9.3  
151.5,41.3,58.5,18.5  
180.8,10.8,58.4,12.9  
8.7,48.9,75,7.2  
57.5,32.8,23.5,11.8  
120.2,19.6,11.6,13.2  
8.6,2.1,1,4.8  
199.8,2.6,21.2,10.6  
66.1,5.8,24.2,8.6
```

Cancel

Submit

Scroll down and press “submit”. Your data points should be shown on the screen like this:

向下滾動並按下「submit」，你的數據點應該會在屏幕上顯示，如下所示：



You can also upload an Excel file to the website directly. Firstly, you will need to prepare your Excel file like this:

你也可以直接將Excel 檔案上傳到網站。首先，你需要準備好Excel 檔案，格式如下：

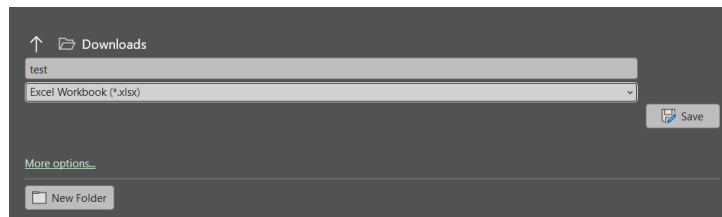
H9			
	A	B	
1	TV,"radio","newspaper","sales"		
2	230.1,37.8,69.2,22.1		
3	44.5,39.3,45.1,10.4		
4	17.2,45.9,69.3,9.3		
5	151.5,41.3,58.5,18.5		
6	180.8,10.8,58.4,12.9		
7	8.7,48.9,75,7.2		
8	57.5,32.8,23.5,11.8		
9	120.2,19.6,11.6,13.2		
10	8.6,2.1,1,4.8		
11	199.8,2.6,21.2,10.6		
12	66.1,5.8,24.2,8.6		
13	214.7,24,4,17.4		
14	23.8,35.1,65.9,9.2		
15	97.5,7.6,7.2,9.7		
16	204.1,32.9,46,19		
17	195.4,47.7,52.9,22.4		
18	67.8,36.6,114,12.5		
19	281.4,39.6,55.8,24.4		
20	69.2,20.5,18.3,11.3		
21	147.3,23.9,19.1,14.6		
22	218.4,27.7,53.4,18		
23	237.4,5.1,23.5,12.5		

Then go back to our website, press the “Browse...” button and look for the file stored on the computer. After uploading, you will see your data points shown on the graph.

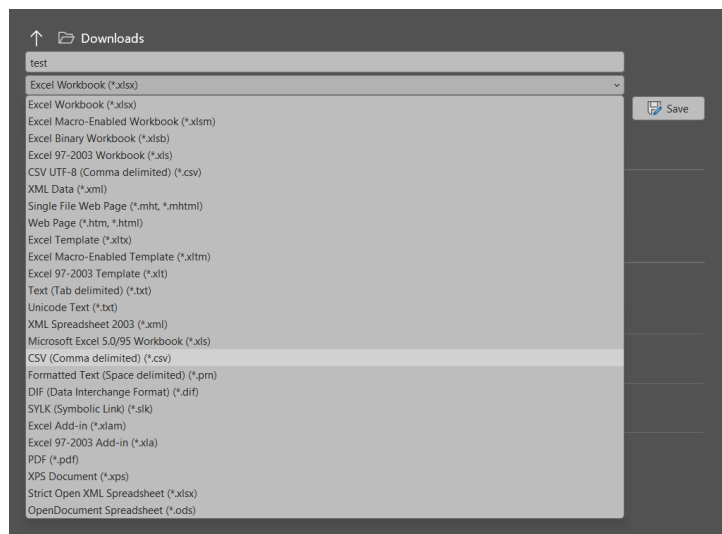
然後返回我們的網站，按下「Browse...」按鈕，尋找儲存在電腦上的檔案。上傳完成後，你會看到數據點顯示在圖表上。

If you face any problem regarding the upload, you may need to save your Excel file in .csv format. To do so, you can go to file → save as, and you can look for the following:

如果在上傳時遇到任何問題，你可能需要將Excel 檔案儲存為.csv 格式。方法是前往「檔案」→「另存新檔」，然後尋找以下選項：



Expand the dropdown list, and look for the .csv format.
展開下拉列表，尋找.csv 格式。



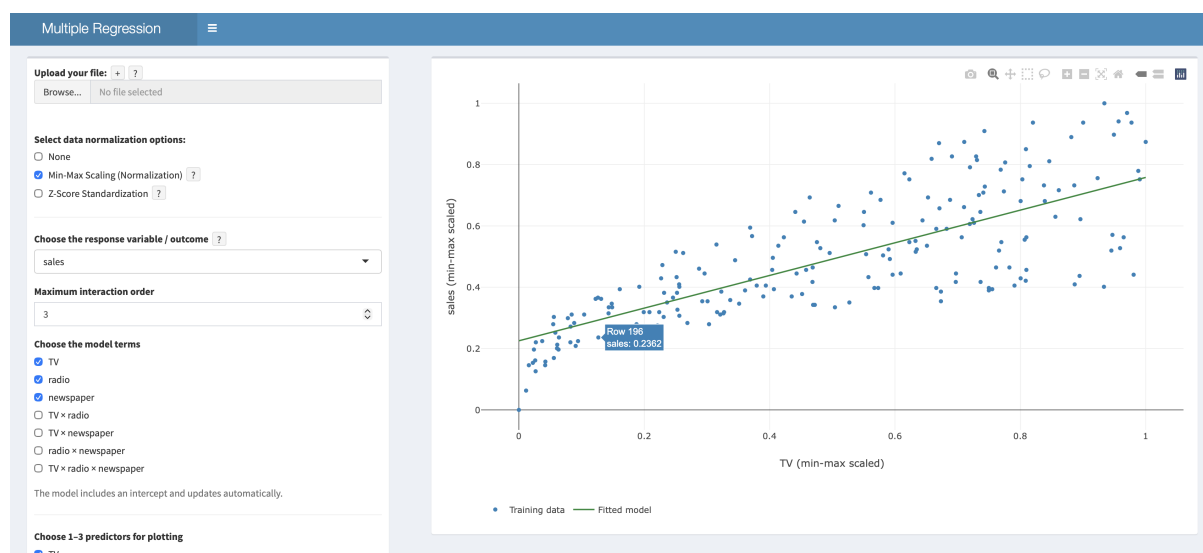
After that, you can upload the file onto our website, and it will work normally.
完成後，你可以將檔案上傳到我們的網站，它就能正常運作了。

2.2 Step 2: Data analysis

步驟二：數據分析

After you have input the dataset, we can now perform multiple linear regression. The following sections will briefly explain the main functions on the page.

輸入數據集後，我們現在可以進行多元線性迴歸。以下部分將簡要介紹頁面上的主要功能。



If you have any questions about the tool, you can always click the boxes with “?” for a short introduction.

如果你對這個工具有任何疑問，可以隨時點擊帶有「？」的方框，那裡會有簡短的介紹。

2.2.1 Choose the Response Variable and the Model Terms

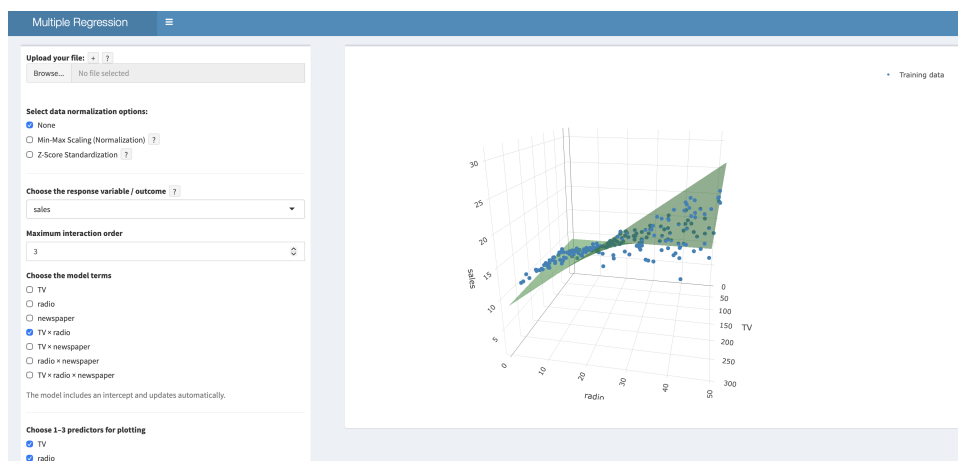
選擇因變量及模型項目

In the section below, you can choose the response variable and the model terms that you want to use for the multiple regression. In our example, we want to model the sales using the advertising expenditure on TV, radio and newspaper. Therefore, we choose **sales** as the response variable, and tick **TV**, **radio** and **newspaper** under “Choose the model terms”. The tool will automatically include an intercept in the model.

在下方的選項中，你可以選擇想要用於多元線性迴歸的因變量及模型項目。在本示例中，我們希望利用電視、電台及報紙的廣告支出來建立銷售額的模型。因此，我們選擇**sales** 作為因變量，並在「Choose the model terms」中勾選**TV**、**radio** 及**newspaper**。工具會自動在模型中加入截距項。

You can also choose interaction terms if necessary. For example, by selecting **TV × radio**, the model will include an additional term describing the interaction between TV and radio advertising expenditure. The maximum interaction order can be adjusted using “Maximum interaction order”.

如有需要，你亦可以選擇交互作用項。例如，勾選**TV × radio** 後，模型會加入一項描述電視及電台廣告支出之間交互作用的項目。「Maximum interaction order」可以用於調整所允許的最高交互作用階數。



For example, if we select only TV, radio and newspaper, the fitted model has the form
 例如，如果我們只選擇TV、radio 及newspaper，擬合模型的形式為

$$sales = \beta_0 + \beta_1(TV) + \beta_2(radio) + \beta_3(newspaper).$$

The values of the coefficients are determined automatically by the fitting process. Therefore, unlike the general fitting tool, we do not need to enter the initial values of the parameters manually.

係數的數值會由擬合過程自動確定。因此，與一般擬合工具不同，我們不需要手動輸入參數的初始估計值。

2.2.2 Data Normalization 數據標準化

The tool also provides three data normalization options: “None”, “Min-Max Scaling (Normalization)” and “Z-Score Standardization”. If necessary, users can select one of the two normalization methods before fitting the model.

此工具亦提供三種數據標準化選項：「None」、「Min-Max Scaling (Normalization)」及「Z-Score Standardization」。如有需要，使用者可以在建立模型前選擇其中一種標準化方法。

“None” leaves the data unchanged. Min-Max Scaling transforms the data to a common range, while Z-Score Standardization rescales the data according to its mean and standard deviation. The normalization option is optional.

「None」表示不對數據作任何變換；Min-Max Scaling 將數據轉換至共同的範圍，而Z-Score Standardization 則根據數據的平均值及標準差重新調整尺度。數據標準化屬於可選功能。

Select data normalization options:

- ☒ None
- ☐ Min-Max Scaling (Normalization) ?
- ☐ Z-Score Standardization ?

2.3 Step 3: Regression analysis

步驟三：迴歸分析

After selecting the response variable and at least one model term, the tool automatically displays the fitted regression model and fitted coefficients.

選擇因變量及至少一個模型項目後，工具會自動顯示擬合迴歸模型及擬合係數。

The “Fitted model” section displays the regression model obtained from the selected terms, while the “Fitted coefficients” section gives the estimated values of the coefficients.

「Fitted model」部分會顯示根據所選模型項目所得的迴歸模型，而「Fitted coefficients」部分則會列出模型中各係數的估計值。

The tool also allows users to choose 1–3 predictors for plotting. This is useful for visualizing the fitted results when several explanatory variables are involved.

工具亦讓使用者選擇1–3個自變量作圖。當模型包含多個自變量時，此功能有助於將擬合結果可視化。

Users can also select “Show prediction on plot” to display the prediction on the graph.

使用者亦可以選擇「Show prediction on plot」，在圖表上顯示預測結果。

2.3.1 Test set evaluation

測試集評估

The tool provides a “Test set evaluation” section for assessing the prediction performance of the fitted model on additional data. To use this function, upload or enter the original (unscaled) test-set values, including the observed outcome.

工具提供「Test set evaluation」部分，用於評估擬合模型對額外數據的預測性能。使用此功能時，請上傳或輸入原始（未標準化）的測試集數據，包括實際觀察所得的結果。

After the test set is loaded, the tool calculates four commonly used error measures: Mean Squared Error (MSE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE).

載入測試集後，工具會計算四種常用的誤差指標：均方誤差(MSE)、平均絕對誤差(MAE)、均方根誤差(RMSE)及平均絕對百分比誤差(MAPE)。

In general, a smaller error value indicates a smaller difference between the observed and predicted values for the test set.

一般而言，較小的誤差值表示測試集中的觀察值與預測值之間的差異較小。

2.4 Worked example: Advertising data

示例：廣告數據

In our example, we use `data_MLR.xls`. If we select `sales` as the response variable and `TV`, `radio` and `newspaper` as the model terms, the fitted model takes the form

對於本示範，我們使用`data_MLR.xls`。如果選擇`sales` 作為因變量，並選擇`TV`、`radio` 及`newspaper` 作為模型項目，模型形式為

$$sales = \beta_0 + \beta_1(TV) + \beta_2(radio) + \beta_3(newspaper) + \epsilon.$$

The fitted coefficients obtained from the dataset are approximately
從此數據集所得的擬合係數約為

$$sales = 2.939 + 0.04576(TV) + 0.1885(radio) - 0.001037(newspaper).$$

This equation gives the fitted value of `sales` for given values of `TV`, `radio` and `newspaper`. It is an example of how the tool converts a dataset into a multiple linear regression model. 對於指定的`TV`、`radio` 及`newspaper` 數值，此方程式可給出`sales` 的擬合值。這示範了工具如何將數據集轉換成多元線性迴歸模型。

3 Conclusion

總結

The Multi-Regression R Shiny tool provides a convenient way for us to apply multiple linear regression in mathematical modelling. Through data normalization, simplified variable selection, visualization and validation functions, we can build and evaluate regression models efficiently.

多元線性迴歸R Shiny 工具提供了一個方便的平台，讓我們應用多元線性迴歸進行數學建模。透過數據標準化、簡化變量選擇、可視化及驗證功能，我們可以有效建立及評估迴歸模型。