

以制限最佳化來解決邊界違反的問題

假設新北市有二十個選區，其投票率與每人平均所得水準高低和藍綠政黨上屆得票率差距兩因素有關，在某次選舉中，這二十個選區的相關資料如下：

選區編號	本屆投票率 (%)	所得 (萬元/月)	上屆藍營得票率 (%)	上屆綠營得票率 (%)
1	43%	4.5	40%	50%
2	80%	3.6	45%	43%
3	66%	5.2	32%	59%
4	48%	6.8	65%	30%
5	76%	4.7	56%	40%
6	57%	5.1	46%	50%
7	81%	3.8	40%	52%
8	79%	5.5	37%	61%
9	58%	4.4	48%	44%
10	64%	5.1	54%	40%
11	69%	5.2	28%	36%
12	58%	4.1	16%	72%
13	61%	4.2	72%	18%
14	45%	5.9	25%	40%
15	84%	3.2	41%	40%
16	62%	5.6	37%	42%
17	54%	4.8	60%	39%
18	59%	5.3	26%	65%
19	71%	4.8	47%	49%
20	75%	4.6	39%	45%

- 令 Y, X_1, X_2 分別代表各區投票率、每人平均所得水準、藍綠政黨上屆得票率差距 (取絕對值)， $\beta_0, \beta_1, \beta_2$ 為迴歸係數，請說明：
 - 線性迴歸模型的數學設定 (請將所得變數 X_1^* 採中心化操作，即 $X_1^* = X_1 - \bar{X}_1$)
 - 各迴歸係數的意義
 - 其「線性」的意義。
- 我們知道採用最小平方方法 (最大概似法亦同) 的迴歸係數解為 $\hat{\beta} = (X'X)^{-1}X'Y$ ，請求出其解，並且解釋其結果。

3. 我們知道投票率的理論預測可由模型推出 ($\hat{Y} = X\hat{\beta}$)，請算出各選區的預測投票率值，並檢驗是否有不合理的情況。
4. 倘若我們知道經驗上 (上表 20 個選區實際資料顯示) X_1^* , X_2 各有其最大最小值，請利用經驗上兩自變量所有出現的可能數值，來討論第二小題所解出的最小平方解是否滿足依變項的值域限制。
5. 請用邏輯上兩自變量所有出現的可能數值，假定各區每人平均所得是在 [1 萬,15 萬] 之間，再重新評估第四題的問題。
6. 請運用老師附件中所準備的 Matlab 制限最佳化程式，評估滿足依變項界域情況下，再考量自變量的經驗和邏輯值域，其各自算出的迴歸係數解為何？
7. 再一次，請用邏輯上兩自變量所有出現的可能數值，即假定各區每人平均所得是在 [1 萬,15 萬] 之間，再重新評估一次預測投票率值，並與第五題的預測結果相比較。

附件：問題的操作化，參數設定，與 **TRMCO** 程式

問題的操作化

首先，我們得將上面的表格轉成 Matlab 可以辨認的格式，或者我們要自行輸入這些資訊，舉例來說，如果我們要輸入前 10 個觀察值的內容，那麼可以這樣做：

```
%% 建立資料矩陣 %%
casen = [1:1:10]';
y = [43;80;66;48;76;57;81;79;58;64];
x0 = ones(length(casen),1);
x1 = [4.5;3.6;5.2;6.8;4.7;5.1;3.8;5.5;4.4;5.1];
x2_kmt = [40;45;32;65;56;46;40;37;48;54];
x2_dpp = [50;43;59;30;40;50;52;61;44;40];
x2 = 0.01*abs(x2_kmt - x2_dpp);
% 這裡則是要把上面做好的y, x0, x1, x2 column合併成一個大的資料矩陣 %
ex7 = [y x0 x1 x2];
% 然後把它存成"ex7.mat"的資料格式，之後Matlab就可以讀取它 %
save ex7.mat ex7
% 這邊的目標則是想要做出中心化的效果 %
x1_avg = mean(x1);
x1c = x1 - x1_avg;
```

至於如何求得回歸係數，舉例來說，如果現在沒有要中心化，那麼我們可以直接用上面的行向量來做出矩陣 X 和矩陣 y ，然後用公式得到 $(X'X)^{-1}X'Y$ 的結果：

```
x = [x0 x1 x2];
beta = (x'*x)\x'*y;
fprintf('beta0:%6.3f, beta1:%6.3f, beta2:%6.3f\n',beta(1),beta(2),beta
    ↪ (3));
```

那麼，如何得到預測值，也當然可以用 Matlab 來算出，這個部分就請大家自行使用了。

參數設定與程式使用

接著，我們在附件中會附有 `ex7_trmco.m` 這個 script，請各位把這個檔案存在自己的資料夾裡，然後在 Matlab 裡把路徑設定到該資料夾，進行自己的參數設定（我在程式的註解當中都寫的很清楚了）。若要執行，則直接在 Matlab 的 Command Window 中鍵入“`ex7_trmco`”，就可以執行了。

另外，如果各位打開附件的 `ex7_trmco.m` 看到許多亂碼，則請到[這個網址](#)來得到原始程式碼。

作業要求

1. 本次作業請大家於 12/27 日 (週三) 上午 10:20 之前，以上傳到 ceiba 的形式交回。
2. 若有問題，一樣請來信或找助教討論。