Practice#1：線性預測：Near Null Space Approach

目的：

承上一個練習，最小平方法是解決線性預測或無解的聯立方程式最常用的方法。最後所得到的 Normal Equations 提供在無解的情況下『最佳』的一組解。最小平方法所提供的『最佳』解，在空間上的意義上洽是『Orthogonal Projection』(勉強翻譯為「垂直投射」)的意思。本練習提供另外一種求解的方式，除讓同學觀摩不同的解題技巧外，也可以藉由比較兩種方法之優劣，練習程式模擬寫作中對於結果的比較。

做程式模擬時需要的資料有時必須自行產生，對於資料的模擬也是本練習的重點。能利用程式模擬出所需要的資料，換句話說，能大量複製資料，才能夠進行

”不同的資料，不同的殘差項，是否對不同的方法產生不同的影響？”

的觀察！

練習：

1. 網頁中提供了三組資料(lp_data1, lp_data2 and lp_data3)，並明示了資料產生的公式，請試著在程式的一開始寫一段小程式去產生這些資料。並練習將這些資料 save 成檔案型態，方便將來可以直接load 出來用。Save 與 load 是 Matlab 處理資料的兩個主要功能。
2. 課堂上講解的 *Near Null Space Method* 其推導過程繁複，不過結果倒是很簡單，也容易寫入程式，問題與其解如下：

Problem: Solve $\tilde{X}^T \tilde{X} \underline{\tilde{a}} = \underline{0}$

Solution:

$$\underline{\tilde{a}}^0 = \frac{\sum_{k=1}^{s} \underline{v}_k(1)\underline{v}_k}{\sum_{k=1}^{s} \underline{v}_k^{\,2}(1)}$$

where $\underline{\tilde{a}}^0 = \begin{bmatrix} 1 \\ \underline{a}^0 \end{bmatrix} = \begin{bmatrix} 1 \\ a_1^0 \\ a_2^0 \\ \vdots \\ a_p^0 \end{bmatrix}$ and $\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_s$ is a set of orthonormal eigenvectors corresponding to zero eigenvalues of the data matrix $\tilde{X}^T\tilde{X}$ 。試著把這個結果寫入程式（即使不瞭解推導過程）。

3. 觀察 data matrix $\tilde{X}^T\tilde{X}$ 的 eigenvalues。Eigenvalue 為 0 的個數與 p 的選擇有關係嗎？試著改變 p 值來進一步觀察。p 值的選擇關係到預測的能力嗎？
4. 對這幾組 data 加入一個殘差項(noise)，再觀察 $\tilde{X}^T\tilde{X}$ 的 eigenvalues，這時候對於 eigenvalue 為 0 的判斷是否有影響？與原先沒有加入殘差時，eigenvalue 的「長相」有什麼不同？給予不同份量的殘差，eigenvalues 值的變化為何？值得觀察。

作業

1. 比較單純的最小平方法與 *Near Null Space Method* 的預測能力，在沒有殘差及不同殘差能量(weightings)時。是不是可以畫一張圖來比較不同程度的殘差 vs.預測能力(以 *Root Mean Square Error 來表示*).
2. 換上 sunspot 資料看看兩者的預測能力。對於不同的資料，是否呈現不一樣的預測能力？值得觀察！