

第五章 結論與建議

模擬資料分析結果顯示，不論設定解釋變數對應其係數有三個非零或者僅一個非零時， $JSWT^+$ 的平均解釋變數被挑選個數與 LASSO 相較之下， $JSWT^+$ 比較符合我們的設定。在挑選出正確解釋變數之正確率比較方面， $JSWT^+$ 的正確率皆大於 LASSO 的正確率。而綜觀每個解釋變數挑選比例，當利用 $JSWT^+$ 變數選取流程做變數選取時，我們設定為零的係數對應其解釋變數被挑選的比例明顯小於 LASSO，即 $JSWT^+$ 傾向挑選到正確解釋變數，且較不傾向挑選到係數為零的解釋變數；LASSO 雖然亦傾向挑選到正確解釋變數，但與 $JSWT^+$ 比較之下，卻容易多選到其係數為零的變數。

攝護腺癌資料分析結果顯示，依照 $JSWT^+$ 變數選取流程做變數選取時，卻出現挑不出變數的狀況，即所有係數被壓縮到零。但我們發現 $JSWT^+$ 最後三個壓縮到零的解釋變數為 X_1 、 X_2 、 X_5 與 LASSO 所挑選出的變數相同。若假設正確解釋變數為 X_1 、 X_2 、 X_5 下，我們認為在挑選 $JSWT^+$ 估計量的參數 b 使得 *SURE* 函數有最小值時，為了使 *SURE* 函數有最小值有時可能會出現過度壓縮係數造成所有係數被壓縮到零，但實際上重要解釋變數在係數路徑圖依舊還是可以顯現出來。所以當出現此種狀況時， $JSWT^+$ 估計量的參數 b 必須向下做校正，即可挑選出重要解釋變數。

關於 $JSWT^+$ 與 LASSO 估計量中的參數選定比較，再次強調 $JSWT^+$ 為一個可以利用資料就直接挑選出其中參數的估計量，即使用 $JSWT^+$ 變數選取流程做變數選取時，我們不需要自訂門檻值或者藉由交叉驗證去找估計量中的理想參數，而估計量的參數可直接經由使得 *SURE* 函數有最小值的方式去找到理想值，所以只會得到唯一的係數路徑。相反的，LASSO 估計量必須先去微調限制式的門檻值 t ，才可能到估計量的參數值。

總之，當考慮解釋變數個數小於樣本個數情況下， $JSWT^+$ 與 LASSO 在變數

選取的比較結果顯示， $JSWT^+$ 表現的比較好，可直接得到估計量的理想的參數，省去了微調參數的動作。

未來研究將考慮當解釋變數個數大於樣本個數時，如何去處理迴歸矩陣轉正交時出現的問題。當 $JSWT^+$ 出現過度壓縮狀況時，如何建立一個合適的校正機制去挑選變數，及利用其他準則去找出合適的 $JSWT^+$ 其參數讓變數選取的效果更佳。最後考慮是否可以將 $JSWT^+$ 應用於羅吉斯迴歸分析進行分類研究。