資料探勘 Project3 報告

資訊三 E44051259 陳冠霖

實作詳細：

一：Hits

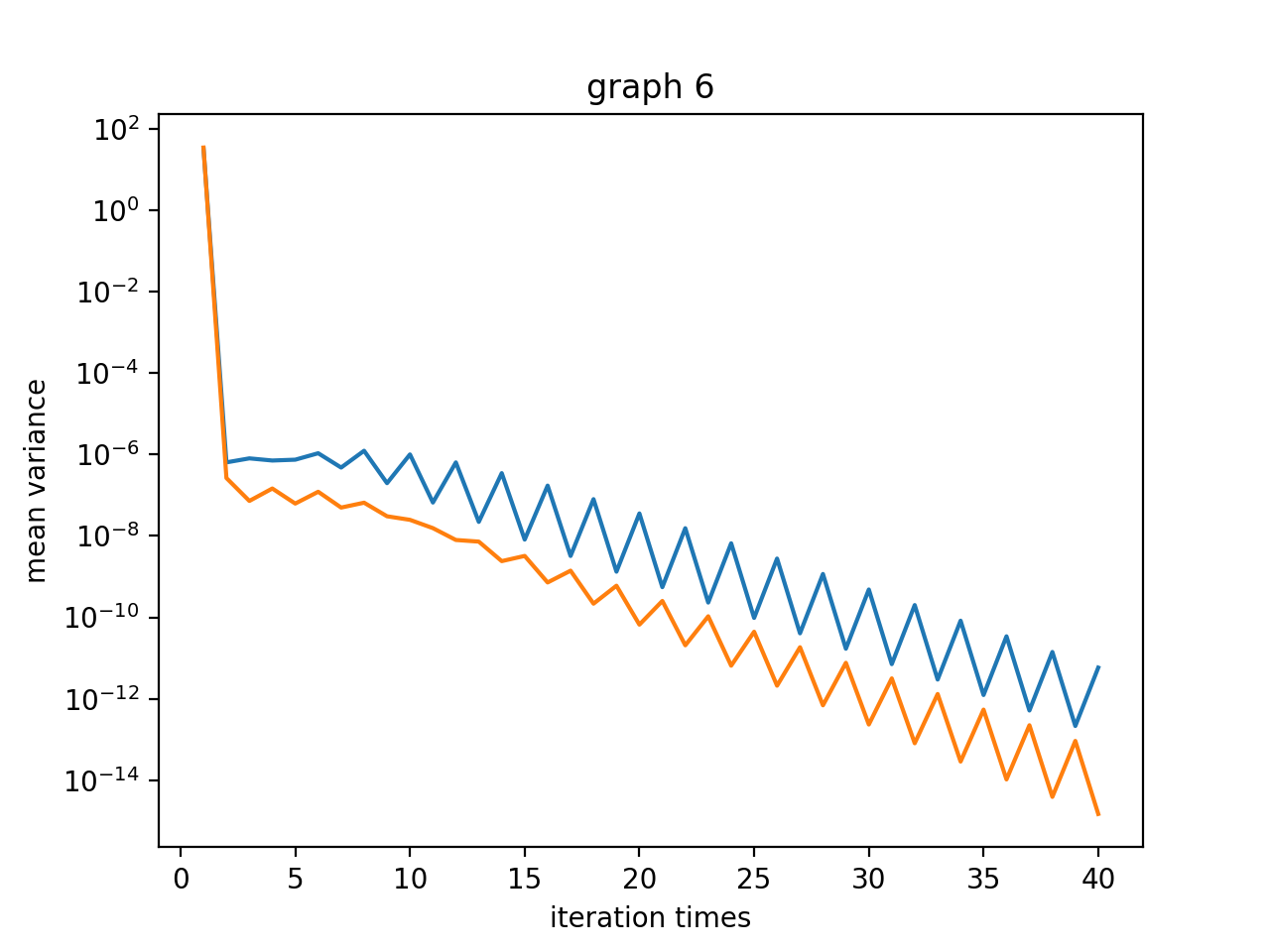
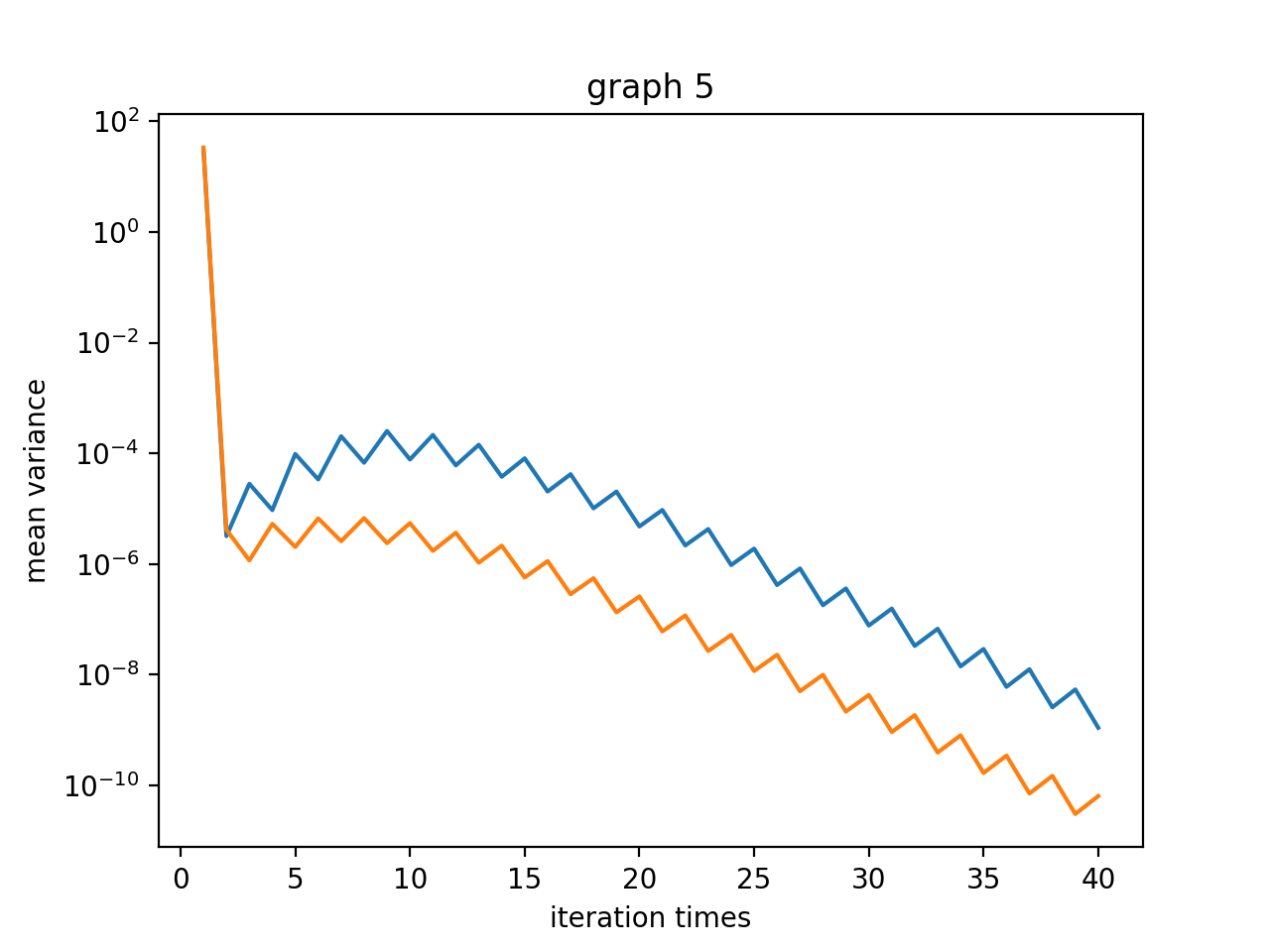
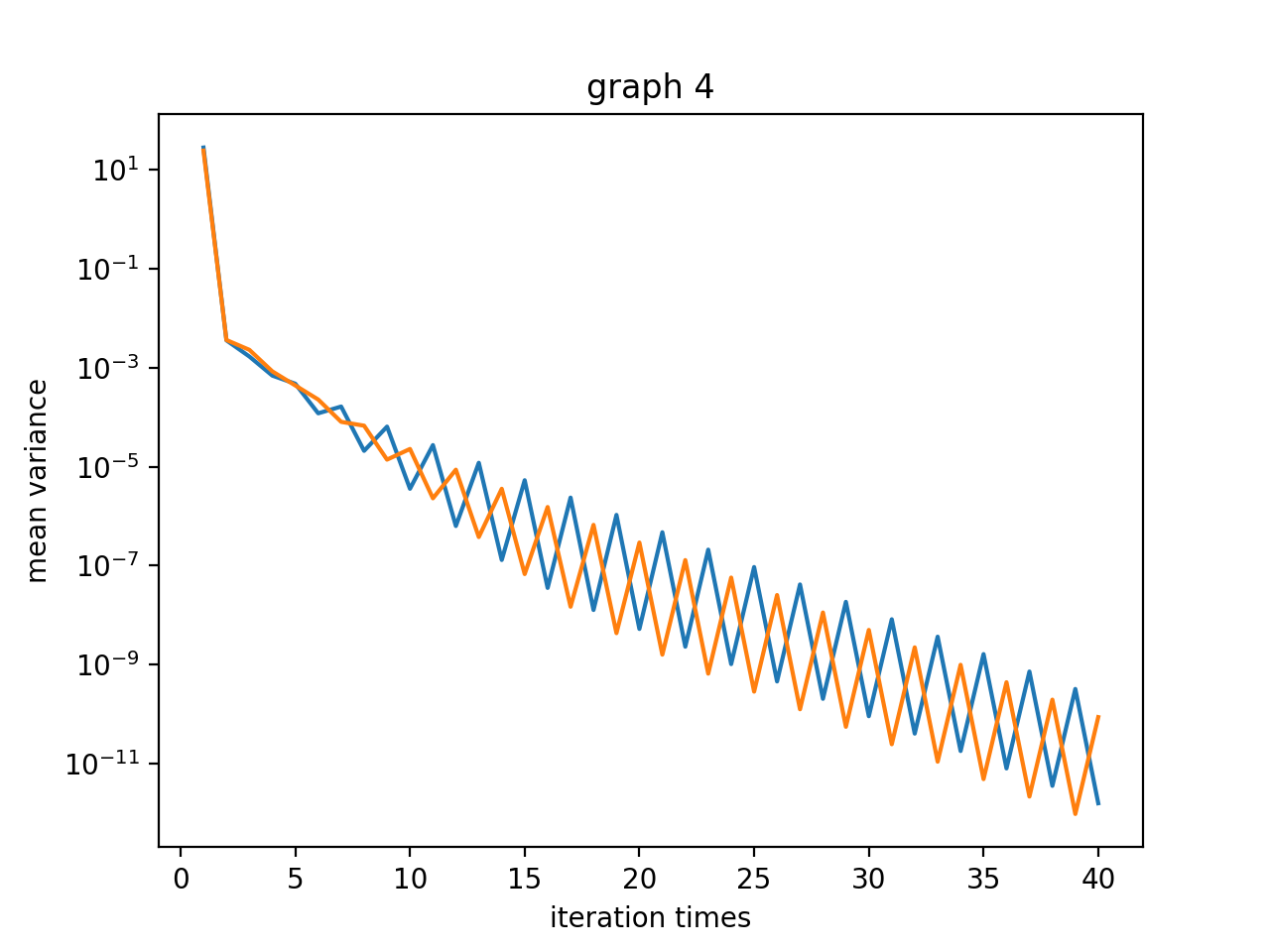
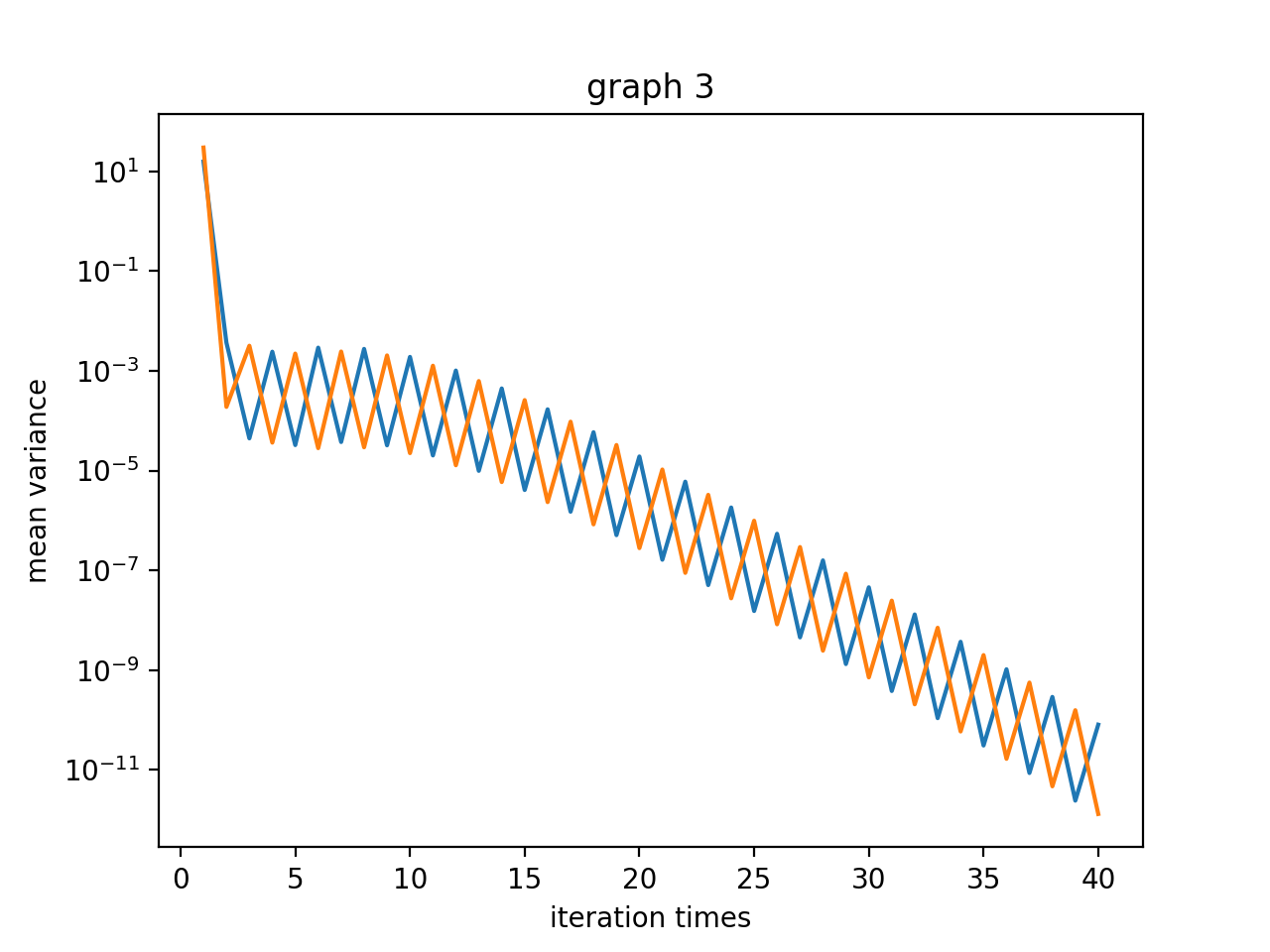
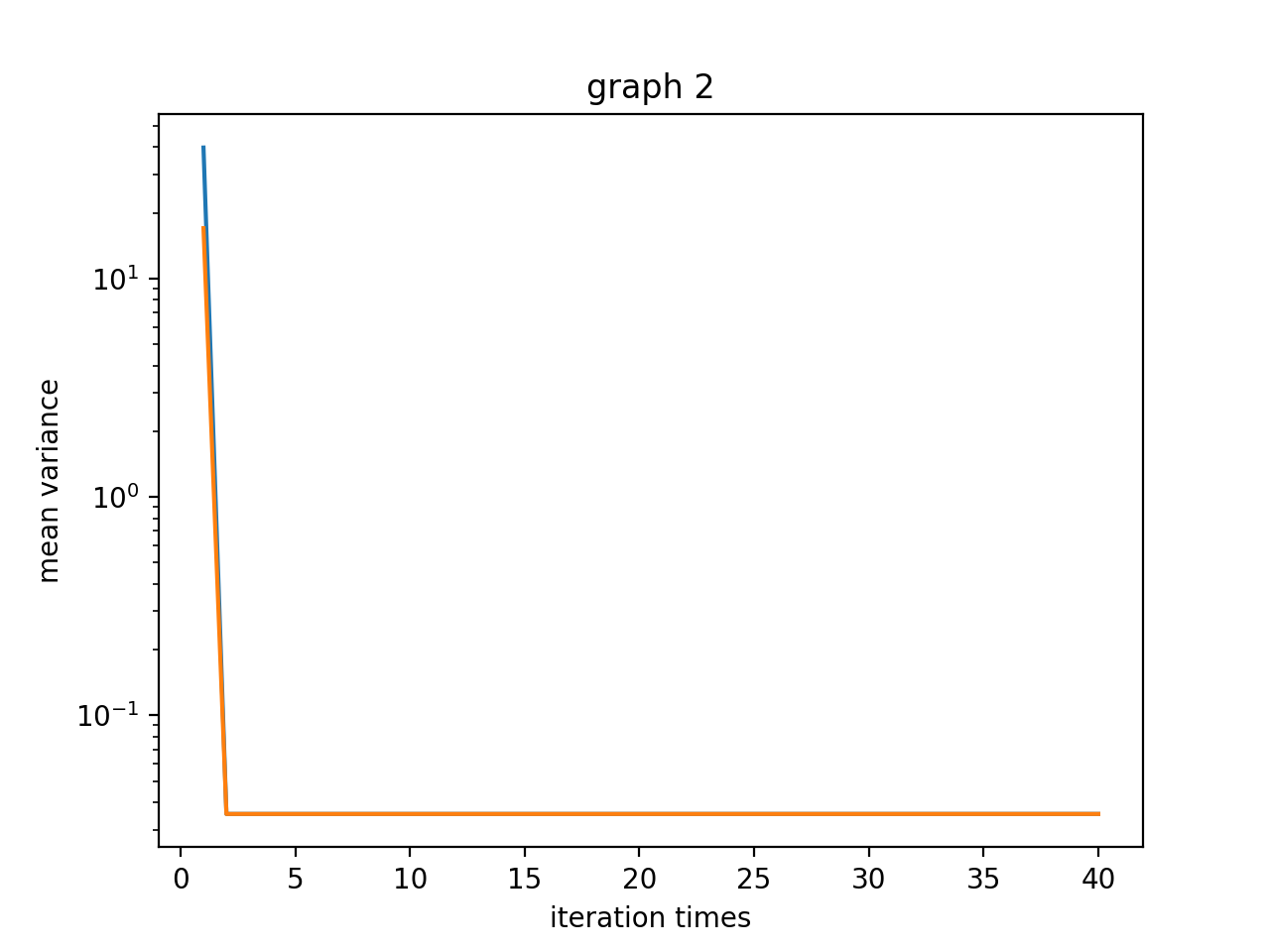
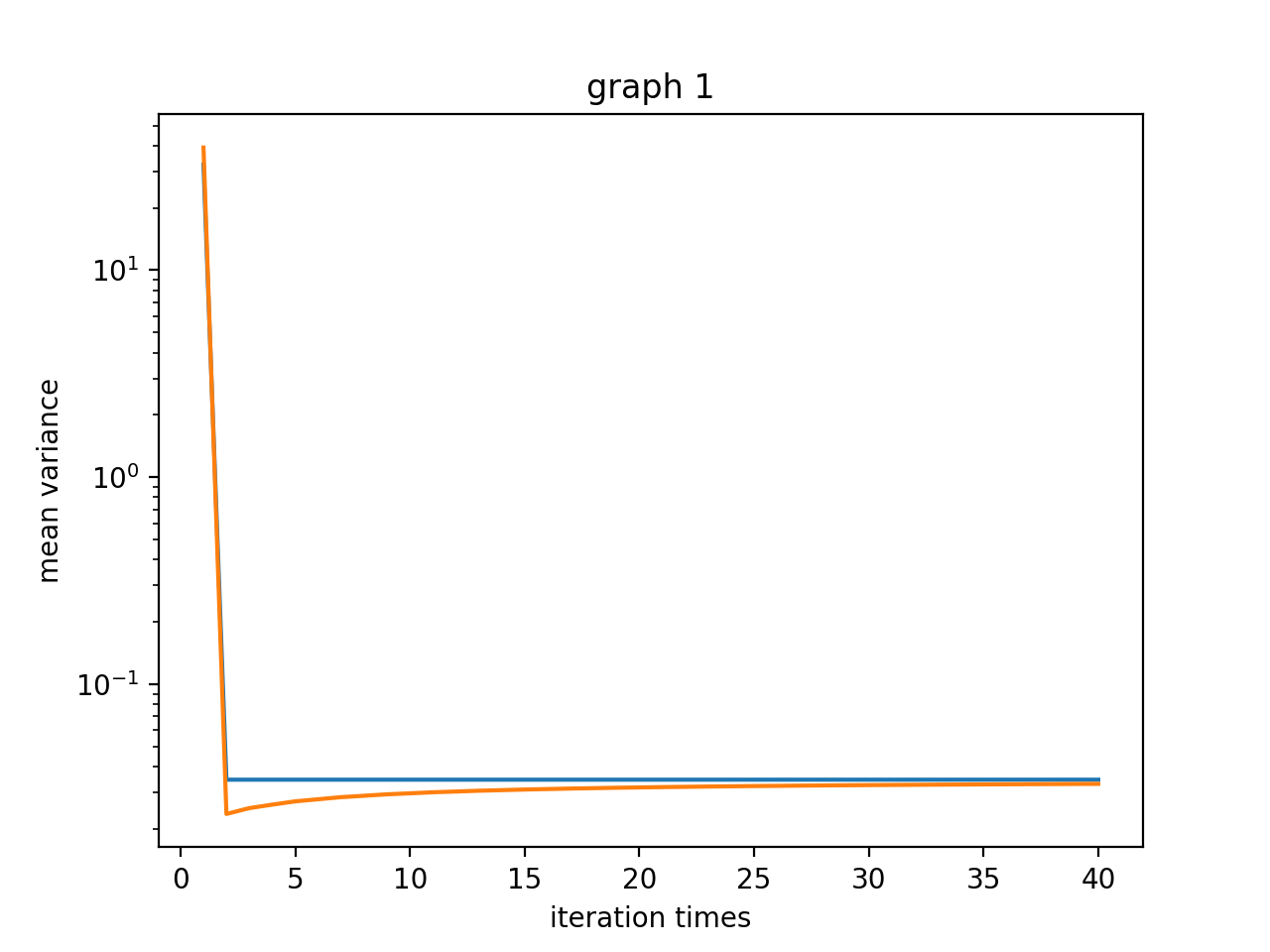
Hits裡面包含兩個數值，一個是hub，一個是authority，而這兩個值，可以隨機給，一開始我是都給1，但這樣在迭代收斂的時候，似乎會怪怪的，所以最後我是給定亂數。而後每次迭代，hub的值為所有child的authority的值的合，而child的authority則為原本的parent的hub值，下去迭代值到收斂。我是設定early\_stopping = 10000，或則收斂的RMSE小於10^-15就停止。

二：

分析與討論：

一：Hits

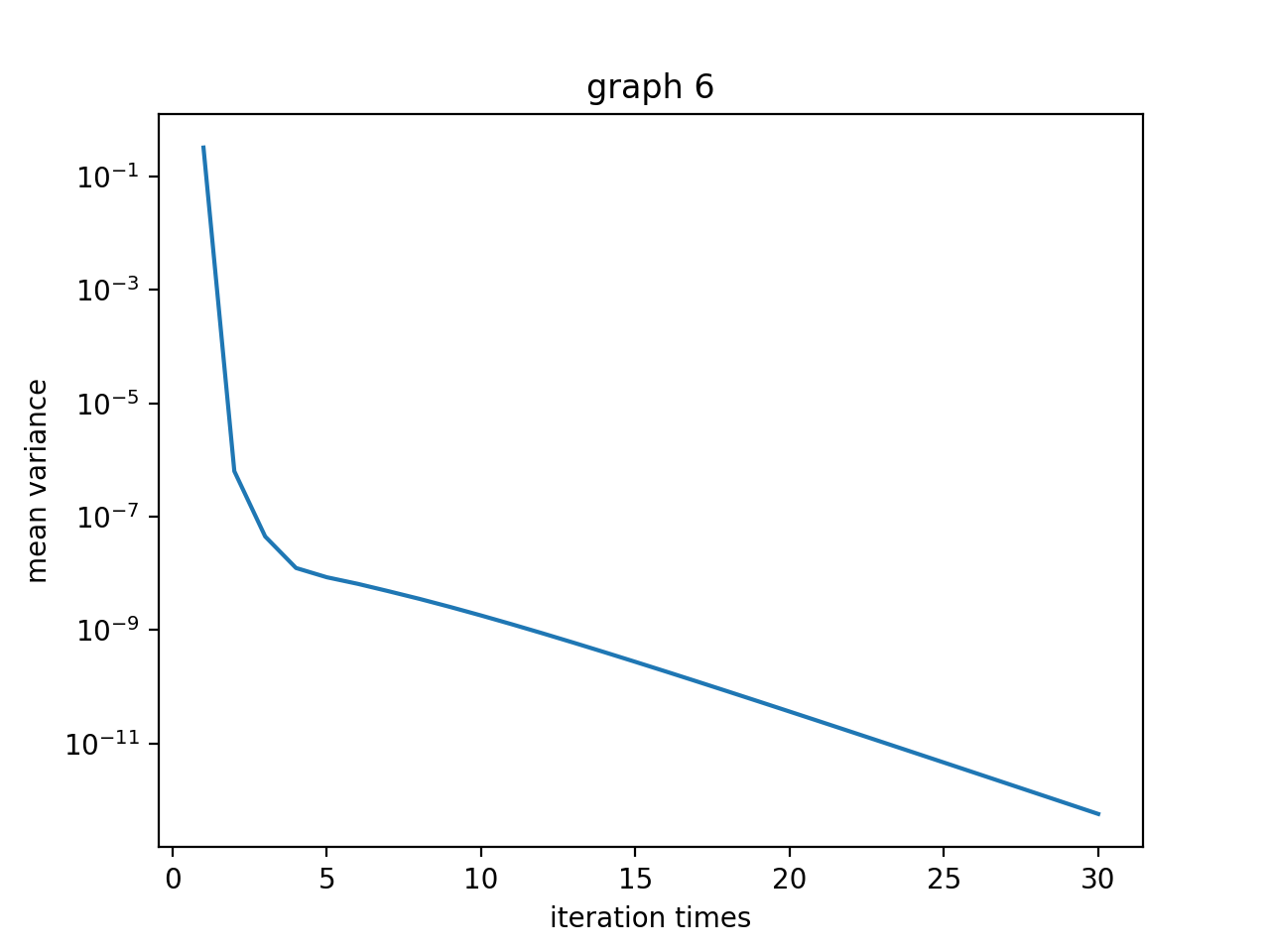
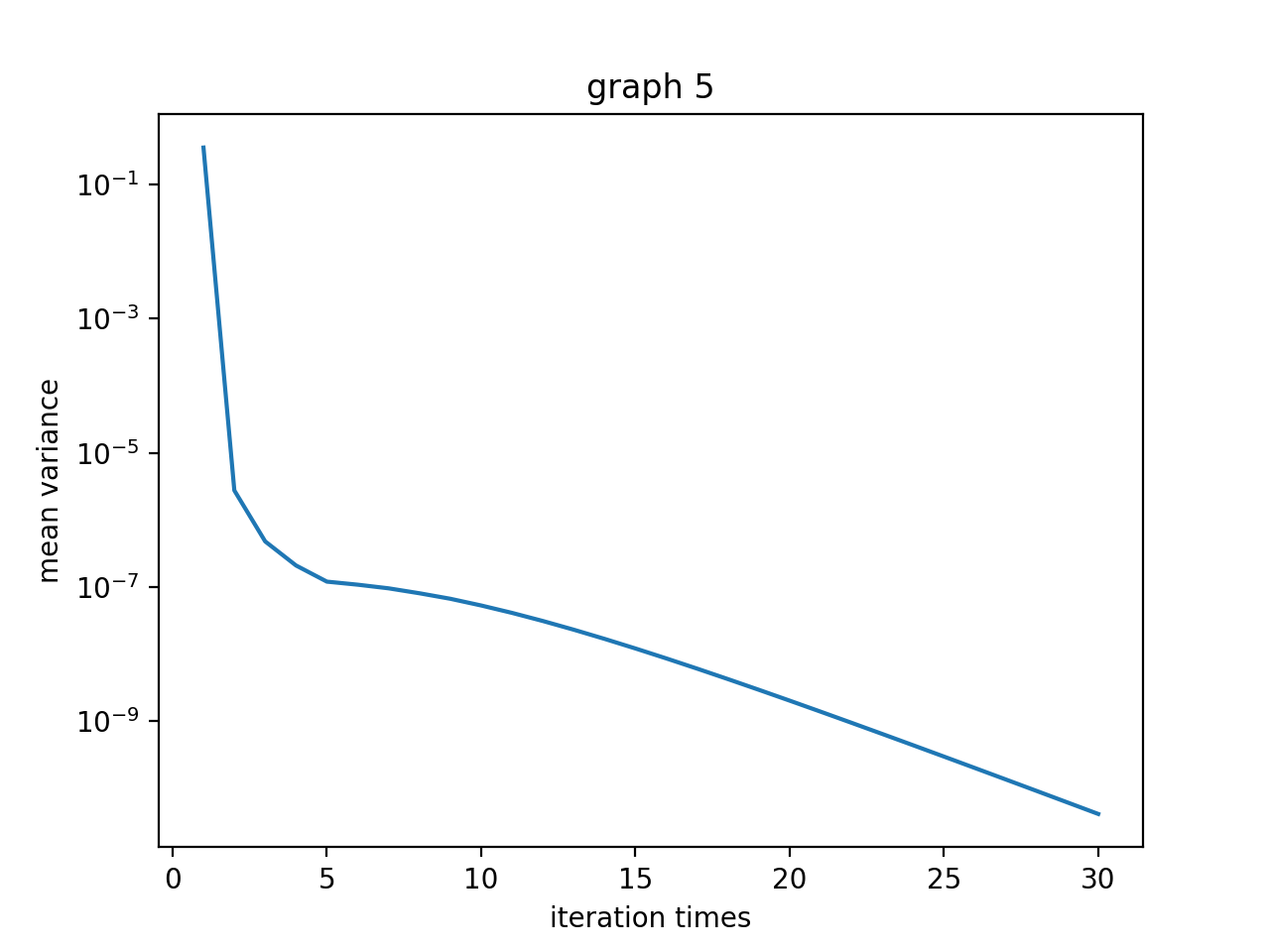
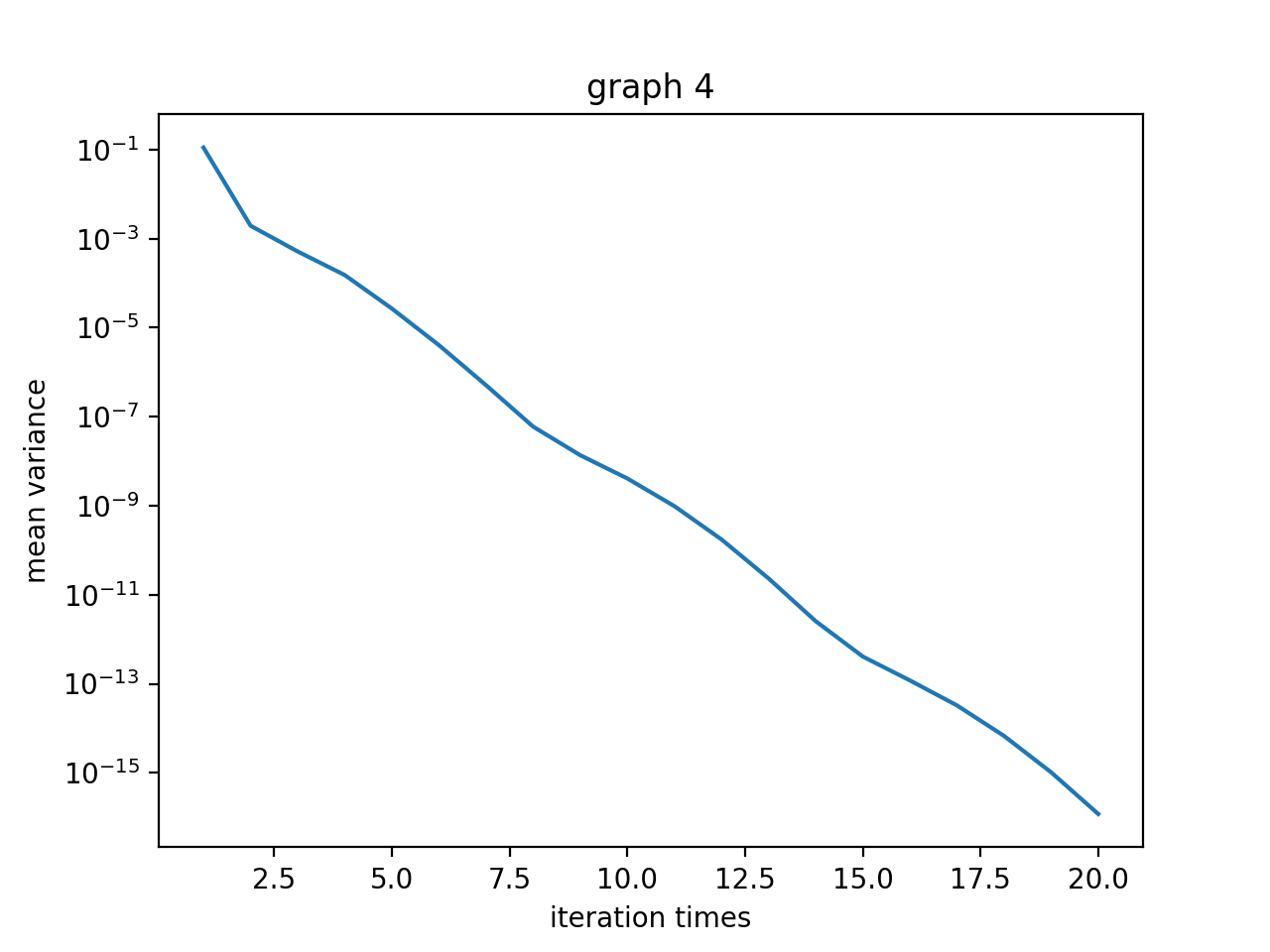
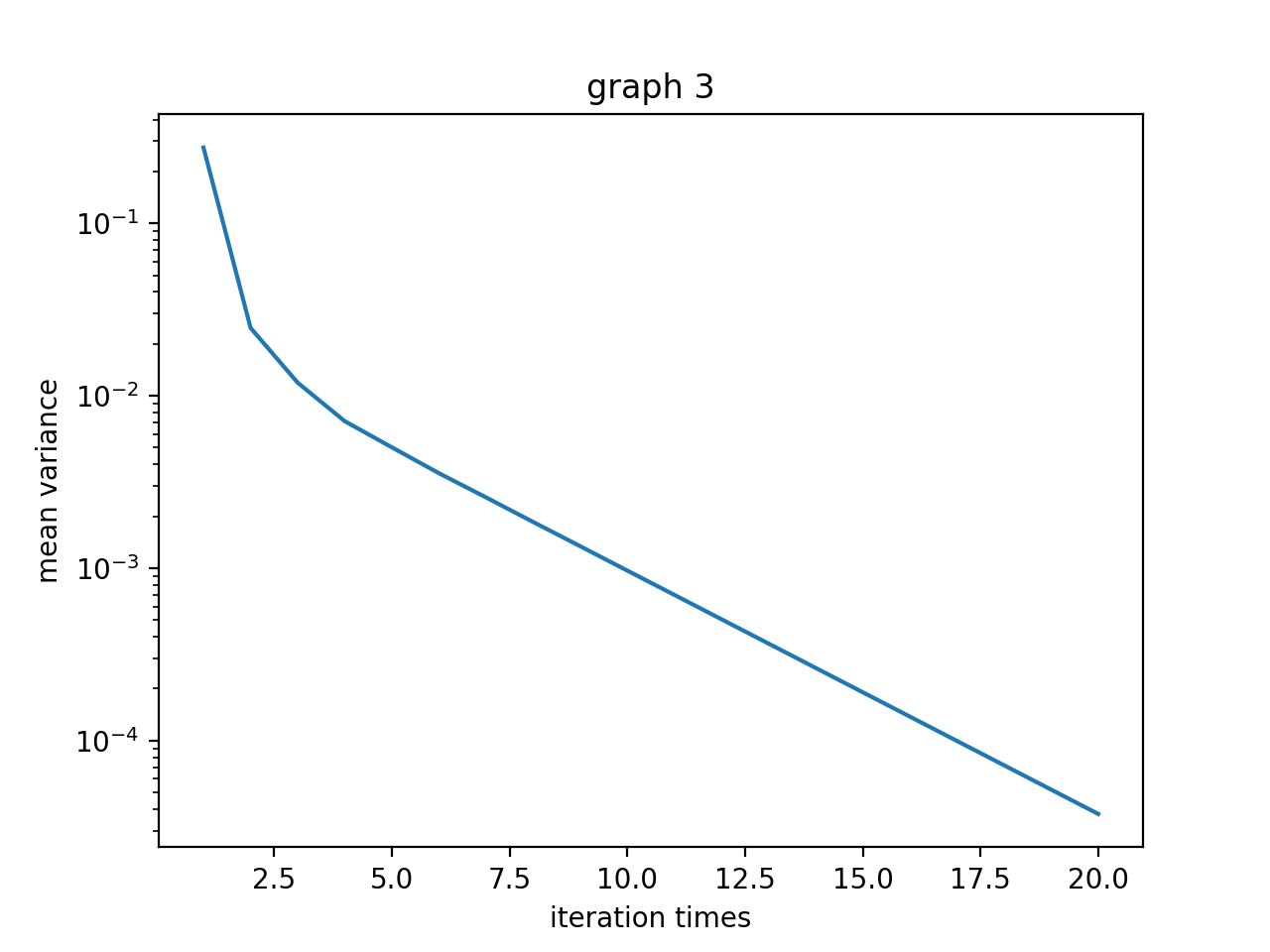
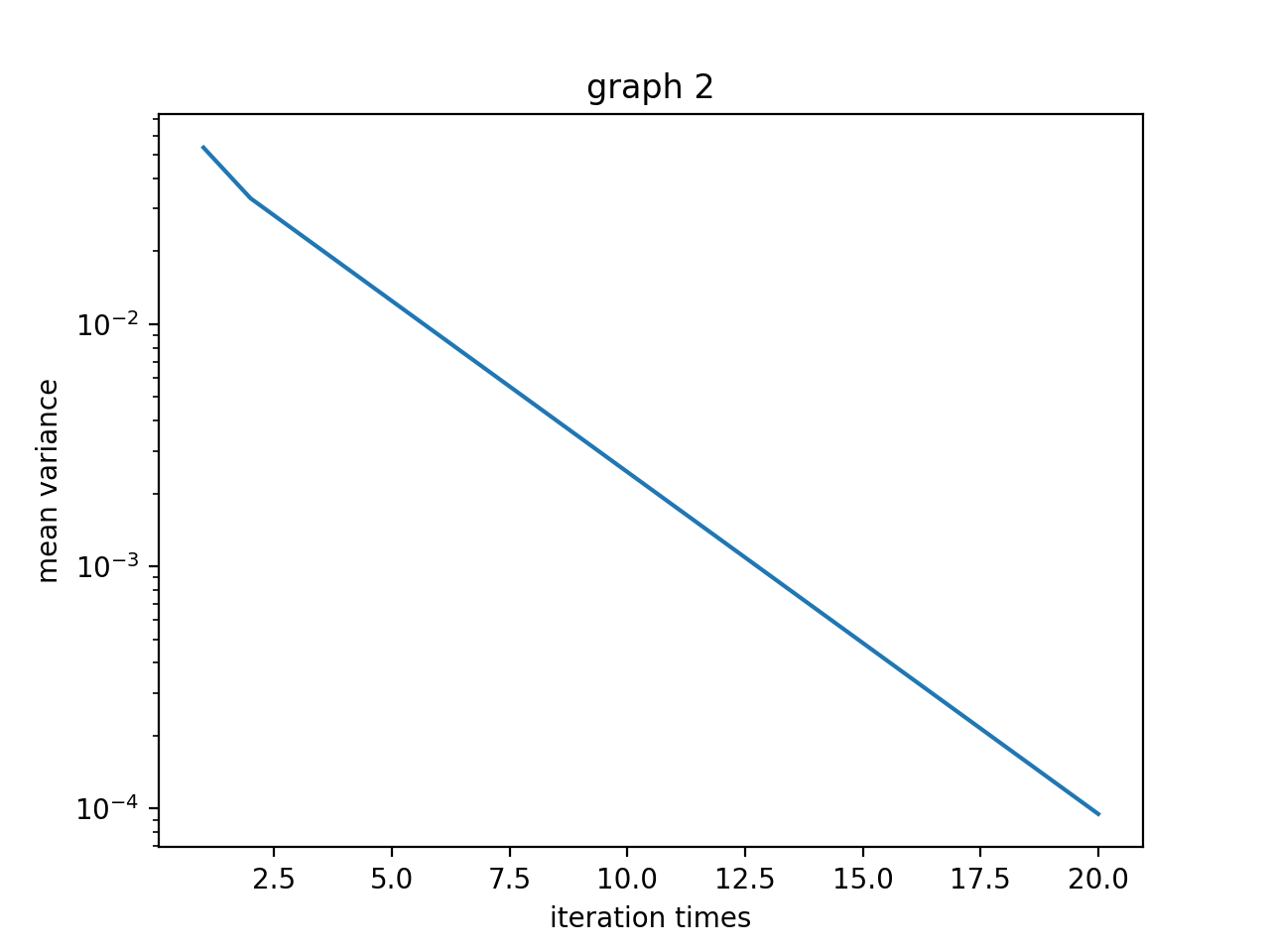
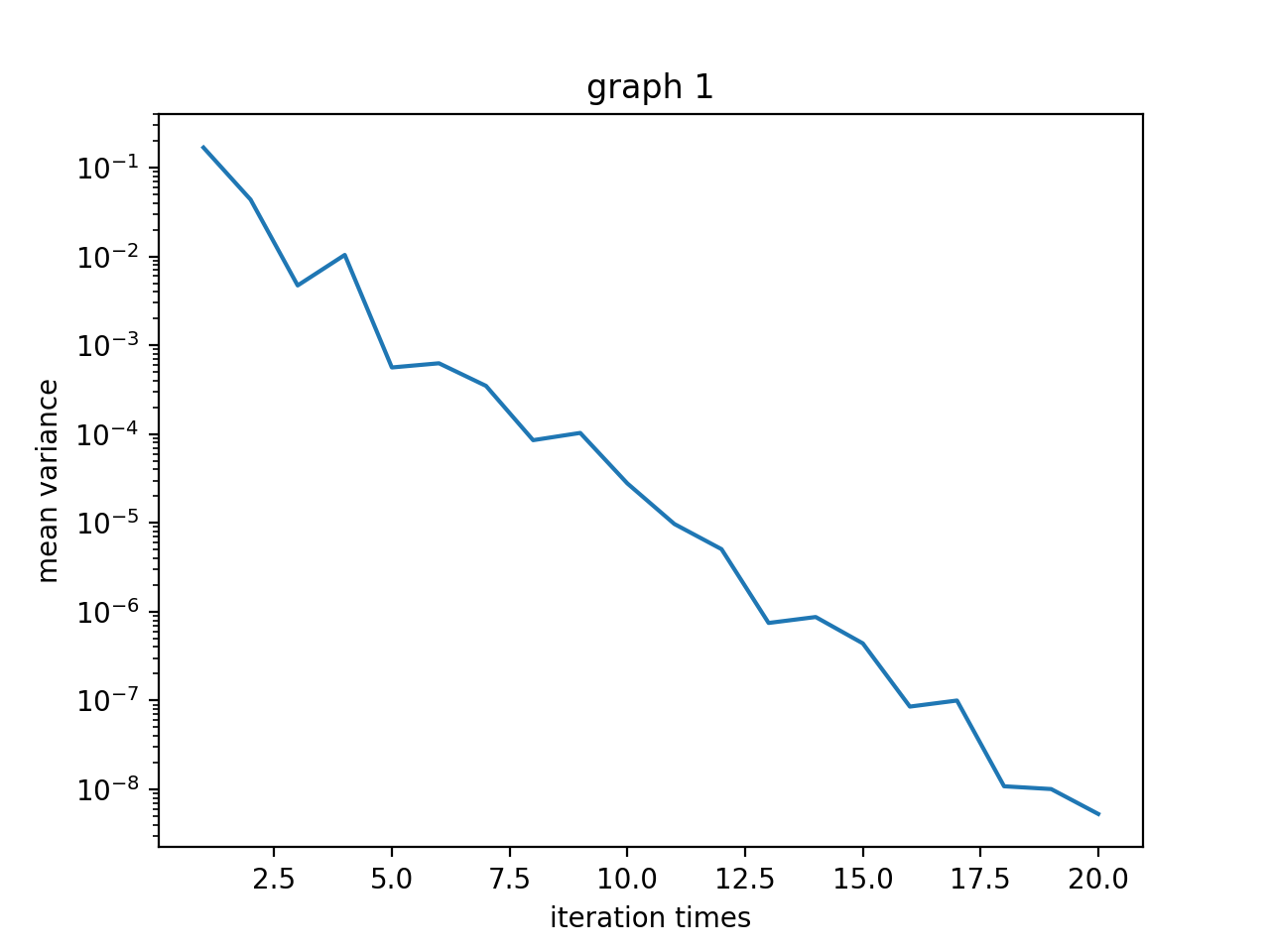
圖片呈現



收斂速度：data1(不收斂)，data2(不收斂)，data3(40次)，data4(68次)，data5(54次)，data6(316次)

二：pageRank

圖片呈現



收斂速度：data1(77次)，data2(203次)，data3(206次)，data4(40次)，data5(145次)，data6(128次)

結論：

1. pageRank的收斂速度比較快