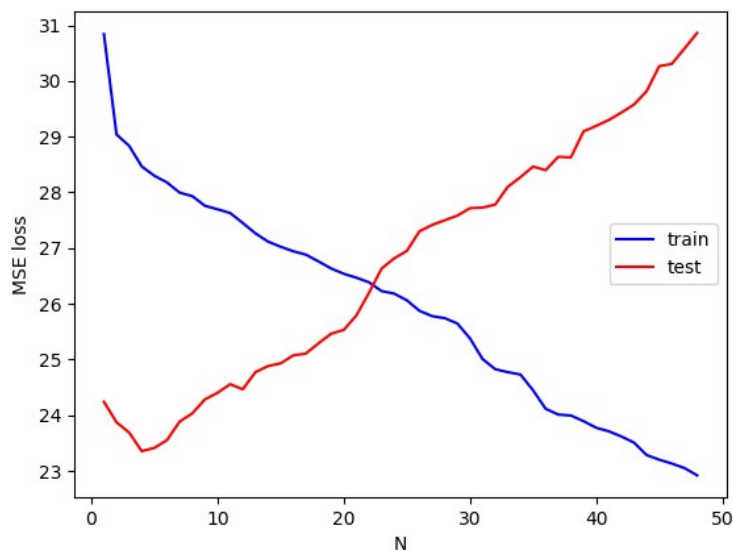
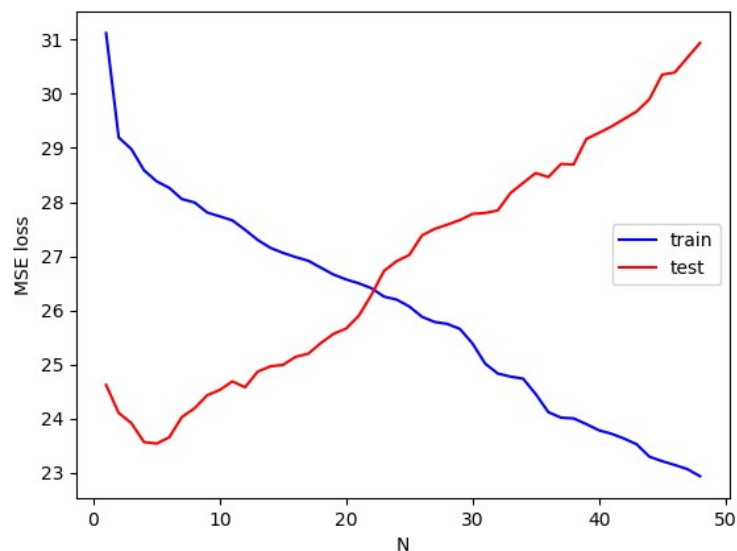


Q2:



在上圖的訓練曲線中可以發現 training error 隨著 N 的增加不斷下降，testing error 在 N 增加時一開始是下降，但到了大約 $N=4$ 的時候就不斷上升。會造成這種現象的原因可能是隨著 N 的增加參數也變多，在 training 時可以更好的 fit 資料，因此 training error 就不斷下降，但伴隨著參數的變多 over-fitting 的現象也越來越嚴重，導致在後來 testing error 便不斷上升。

Q3:



上圖為把 w_0 拿掉後的訓練曲線，可以發現和有 w_0 的情形相差不多。把 w_0 拿掉等同於是把 bias 拿掉，可以達到降低參數的效果，降低 over-fitting 的效應，但是因為 linear regression 相當於只有一層的 neuron network，把 bias 拿掉後對整體參數量的影響不大，故看不出太大的效果。

另外一個可以降低 over-fitting 效應的可行方法是 early stopping，我們可以把 training data 額外分出一部份當作 validation data，慢慢地增加 N 的大小，然後每次都計算 training error 和 validation error，當 N 增加但 validation error 不再下降時就停止增加 N，把當前的 weight 記錄下來，如此一來就比較能停在上圖兩張圖中 testing error 的最低點，避免 over-fitting。經過我實測的結果，我將 training data 中的 3000 筆當作 training data，2688 筆當作 validation data，得到了下圖的訓練曲線。由下圖可發現我們在大約 N=2 的時候便會 early stopping，如此一來 testing error 便不會太大，但此時 testing error 為 24 點多，比起上面用全部 training data 得到的 23 點多的 loss 來得高，可行的解決方法是使用 2-fold cross validation 加上 ensemble：將當作 validation data 的 2688 筆資料當作 training data，原本當作 training data 的 3000 筆資料當作 validation data 再重 train 一次，最後再將兩次得到的 weight 做平均，如此便可再將 testing error 再次降低（實測 testing error 降到了 23 點多）。

