

EDGE INTERNSHIP

作業一 建構回測模型的基礎 劃出績效曲線

李騏維

2025.07.10

一、在每個交易日，皆存在著不同到期日的合約。在你分析的商品中，哪些到期月份的合約的交易量是最大的？

在 TEJ Pro 的資料庫中，可以找到台指期的交易資料，代號 TX。將其以 csv 檔案下載下來，可以得到單筆交易資料的其中以交易日為篩選條件，可以看到期貨的代號、名稱、交易日期、到期月等相關資訊。以交易日期作資料排序，可以區分出每年每月的具體交易。其中交易日期的部分會有重疊，代表同一個交易日中，購買不同期限合約的具體資料。將每筆資料的成交張數(量)再做一次由大至小的排序，可以得到每筆資料中的各項數據。將所有交易資料的合約整理後，發現每筆合約的共同特性：在每一筆合約中，到期日越近，其成交張數越多，而最快到期的合約便是於購買後下個月便到期。故可整理出以下資訊：

於交易當天的下個月到期的合約中，會有最大的交易量。

將該種類的合約稱為近期合約 C_n 。於此同時，另外定義有次大交易量的合約做為遠期合約 C_f ，以便後續分析之用。

二、市場有慣性交易的熱門月合約。請以最接近到期的熱門月合約作為持倉依據。

此即為方才在第一題定義的合約 C_n 。

三、請製作出一個函數，能給定在到期日前 N 日進行轉倉(rolling)，以收盤價計算出每日損益(daily return)。

首先以轉倉的定義起手：對於即將到期的合約，投資人可在到期日之前將合約賣出，轉而購買到期期限更晚的合約，其中轉倉的花費作為轉倉成本，會一併於下題做討論。在此先列出求取轉倉天數，到期日前 N 日的求解邏輯如下(其中交易期間取十年，自 2014 到 2024 作分析)：

1. 找出最大與次大的交易量做為主力與次主力，並合併在一個檔案中
2. 用代號(Eg. TX201401、TX201402 做排序)，【剩餘天數】亦會做好排序
3. 設定轉倉日 N，對想要設定的 N 分別做一次計算
4. 在【剩餘天數】欄位，若剩餘天數>N 時，採用該日的報酬率(【報酬率】欄位)

5. 若剩餘天數=N(或<N 的第一天)時，做轉倉
6. 轉倉的概念是:找到【下個代號】的【同一天日期】，換成下個代號的合約，如此往復
7. 做累積報酬與其他統計

四、請注意在轉倉時會有相對應的轉倉成本，數值可能為正或負。思考該如何正確地處理它以得到正確的損益曲線。

對於轉倉成本的正負，我認為應該由多空市場的角度切入，理由是多空市場的不同會直接影響到轉倉成本的正負。首先定義轉倉成本 C_{roll} ：

$$C_{roll} = |\text{近月收盤價} - \text{遠月開盤價}|$$

事先定義轉倉成本恆正，再藉由多空趨勢決定其正負。而在每一筆交易下都會有其相對應的多空趨勢，將每個交易的作多做空選擇的 PnL 質分別算出並建立簡單的迴圈迭代，並賦予該時段的交易一個多空變數 $best_dir$:

if PnL_{long} > PnL_{best} , best_dir = 'long'

elif PnL_{short} > PnL_{best} , best_dir = 'short'

其中 PnL_{best} 經由迭代獲得。至此便可以決定 C_{roll} 的正負:

if best_dir == 'long' , C_{roll} = C_{roll}

elif best_dir == 'short' , C_{roll} = -C_{roll}

得到 C_{roll} 的正負後，便可以對問題三中討論的模型進行優化：

1. 出最大與次大的交易量做為主力與次主力，並合併在一個檔案中
2. 用代號(Eg. TX201401、TX201402 做排序)，【剩餘天數】亦會做好排序
3. 設定轉倉日 N，對想要設定的 N 分別做一次計算
4. 在【剩餘天數】欄位，若剩餘天數>N 時，採用該日的報酬率(【報酬率】欄位)
5. 若剩餘天數=N(或<N 的第一天)時，做轉倉，**同時考量轉倉成本**
6. 轉倉的概念是:找到【下個代號】的【同一天日期】，換成下個代號的合約，如此往復
7. 做累積報酬與其他統計

值得注意的是，嚴格意義上來說，本模型設定的最大和次大合約並無絕對的區分，亦即兩種合約的角色會隨著時間前進而交錯成為最大與次大，因此若以絕對的【最大交易量】與【次大交易量】來事先區分兩種合約並不妥當，故需要將轉倉的資料做一次有系統的排序。

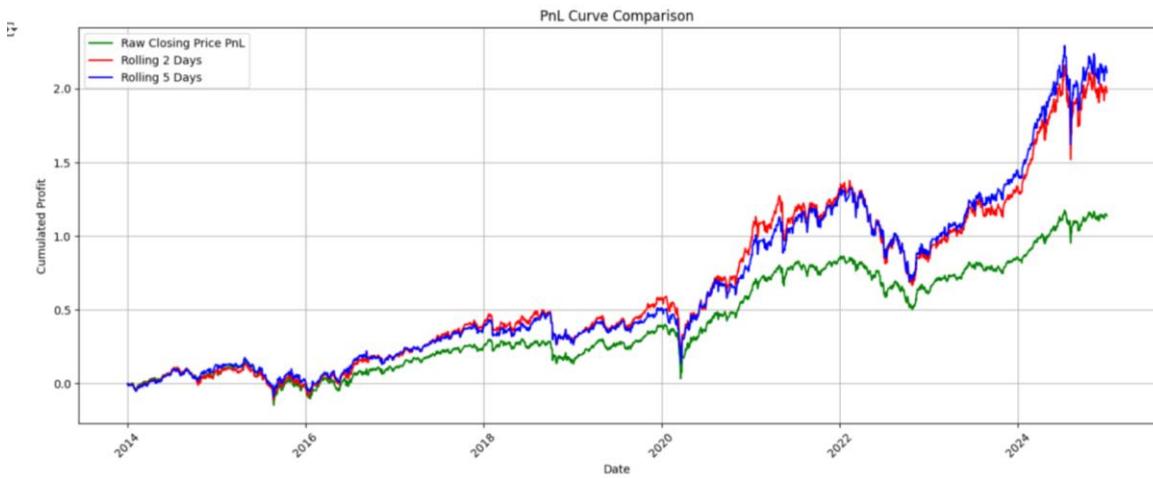
轉倉的概念與雙股螺旋的概念類似，輔以簡易示意圖如下：

(橫軸為交易量，縱軸為時間)



五、最後請將每日損益進行累積並作圖，將其與單純連續價格計算出的 PnL 進行比較：請問你處理的商品的轉倉成本是偏正還是偏負？是什麼原因影響了轉倉成本？

透過轉倉一行為，我得到了比以每日收盤價做交割的累積報酬更高的結果，因此可以確定：**透過模擬轉倉的行為，會單純比用每日連續價格去做台指期的買賣，其結果更好。**



六、計算出這段時間內的 Sharpe ratio, annual volatility, annual return, max drawdown, max recovery days，以表格呈現出來

結果表格如下：

	Annual Return	Annual Volatility	Sharpe Ratio	Max Drawdown	Max Recovery Days	
Raw Close PnL	0.112498	0.169225	0.630112	0.315079	508.0	
Rolling 2 Days	0.121913	0.159847	0.719819	0.299733	478.0	
Rolling 5 Days	0.125976	0.156314	0.759228	0.270352	423.0	

在表格最下方，我另外加上了兩欄，分別代表該年度數據期間的轉倉週期與頻率。

七、Python 原始檔與資料來源

Python 原始檔案: <https://colab.research.google.com/drive/1AezhMI4NQkLAIv47-7h08xZO7Po0oga#scrollTo=RMnYvLp-htYL>

台指期月交易資料 : TEJ Pro

台指期每日詳細交易資料 : 臺灣期貨交易所