

## 線上決策與賽局理論

呂及人（資訊科學所）

在日常生活中，我們常面臨的一個困境，就是必須在未知後果的情況下做決定，而後為我們所做的決定付出代價。我們希望避免的是遺憾：後悔我們早知道應該作出不同的決定。如果我們之前沒有遇過類似的情況，那麼我們很難希望能夠保證作出好的決定。然而若是我們常在類似的狀況下作決定，那麼我們就有希望能夠從過去的經驗中學習，從而在後來作出越來越好的決定。日常生活中不乏這樣的例子，例如預測天氣、交易股票、或選擇上下班的路等。在資訊科學中，也常遇到這樣類型的問題，包括電腦資源的分配、網路封包路徑的選擇、與網站廣告的刊登等。

從這些情境中，我們可以抽象出如下一個稱為線上決策的問題。考慮一個T回合的遊戲，玩家在每回合中必須採取一個決策，而後由此回合的損失（或獎勵）函數，決定該決策對應的損失（或獎勵），而玩家據此可以調整下一回合的決策方式。我們希望有好的線上演算法（online algorithm），來幫玩家在每回合挑選好的決策。但是什麼是好的決策？我們想要達到什麼樣的目標呢？

一個自然的目標是降低線上演算法的總損失，不過由此很難評估線上演算法的好壞。而另一個常用的評估標準，乃是將其與某類型的離線演算法（offline algorithm）比較。這類型的離線演算法可以在看到所有的損失函數後才作決策，但是被限制必須在所有回合都採取相同的決策。線上演算法與此類最好離線演算法所得總損失的差，被稱為線上演算法的遺憾程度（regret），而降低此遺憾程度乃是線上演算法的重要目標之一。此領域的一個重要結果，乃是設計出了有效率的線上演算法，可以達到約T的平方根這麼低的遺憾程度。而這樣的成果，不只在機器學習方面有深遠的影響，在其他領域，包括演算法設計、複雜度理論、分散式計算等，甚至在其他學科，包括經濟學、優化學、生物學等，也有令人意想不到的應用。

有鑑於上述低遺憾程度演算法的重要性，我們想要更進一步的改進與推廣他們。我們注意到前人的工作，大多考慮的是必須面對最惡劣情境之下最不利的損失函數。然而我們相信，我們周遭的環境並非總是惡意的，而損失函數或許有規律可循。例如，天氣狀況或股價在某一時刻與下一時刻一般都會有所關聯，所以其差異通常不大。為了刻化這種規律，我們定義了一個衡量損失函數偏離總量的參數。我們也設計出如何利用這種環境規律的新演算法，其遺憾程度可以低到約是損失函數偏離總量的平方根。因此，當我們所處的環境以相對穩定的方式演進，其產生的損失函數有較低的偏離總量時，我們的演算法可以達到遠低於前人演算法的遺憾程度。另一方面，當環境確實惡劣，而損失函數的偏離總量達到其最大值T時，我們的演算法仍然得到相同於前人，約T平方根的遺憾程度。因此，我們可以將前人的成果看成是我們成果的一個特例。此外，我們也將成果擴展到更一般性的線上優化問題，並得到類似的結果。

除了設計更好的線上演算法，我們也在其他領域為其尋找新的應用，而其中的一個應用是在賽局理論方面。賽局理論一個重要的研究方向，乃是探討一群為自己利益打算的個體，在彼此利益衝突的情形下會達到什麼樣的可能狀態。我們感興趣的是多回合的賽局，而納許均衡（Nash equilibrium）是一種被廣泛採用來預測這種賽局會達到的可能狀態，因為它是一種一旦到達就不會離開的穩定狀態。然而這也產生了賽局如何達到這樣均衡狀態的問題。事實上，現在資訊科學家普遍認為並不存在有效率的演算法，可以為任何給定的賽局找出一個納許均衡點。這意味著均衡點一般而言可能無法在合理的時間內達到，而平常我們觀察到的狀態，其實可能是遠離任何均衡狀態的。如果是這樣，那麼過去許多基於均衡點的研究，有可能就喪失其意義了。為了解決這種窘境，一個新興的研究方向，乃是探討對於哪一類型的賽局，當參與者採取何種合理的演算法時，則整個賽局系統能夠快速趨於均衡點。我們認為一個參與這種多回合賽局的自私個體，可以被視為面對先前所討論的線上決策問題。而一個對自私個體的合理誘因，乃是降低他在這個多回合賽局中的遺憾程度，因此他有意願採行一個保證低遺憾程度的線上演算法。我們證明了對於某一大類統稱為壅塞性的賽局（congestion game），當每個參與者都採行某一大類低遺憾程度的演算法時，則整個賽局系統確實會很快趨近於某種納許均衡點。此外我們也證明他們所趨近的均衡點，會呈現良好的社會福利狀態（social welfare），其達到的社會福利值，非常接近最好可能的社會福利值。

壅塞性賽局包含許多常見的賽局，例如資源競爭的賽局。這類的賽局，在沒有好的規範下，常常會因為自私的個體採取自私的行為，而導致整個社會陷於不利的均衡狀態。我們的研究顯示，如果我們不要完全放任各個參與者，而是建議他們採用某類型的低遺憾演算法，則他們基於自己的利益仍然有意願採用這樣的建議，而這會導致整個社會很快的趨近一個具有良好社會福利的均衡點。