

野球への動的計画アプローチ

吉良 知文, 稲川 敬介

野球をマルコフゲームとしてモデル化すると、攻撃側にとって勝つ確率を最大にするのは打撃か盗塁かあるいは犠打か、守備側にとって、打者を敬遠すべきか否かといった最適な意思決定を状況別に計算することができる。本稿では、野球をモデル化し、それを解く動的計画アプローチを解説する。約 645 万状態の有限マルコフゲームとして定式化し、状況別の両チームの最善手であるマルコフ完全均衡点とそのときの均衡勝率であるゲームの値を 2 秒未満で計算できることを紹介する。後攻チームの優位性についても触れる。

キーワード：マルコフゲーム，マルコフ完全均衡点，勝率最大化，後攻チームの優位性，最適打順

1. はじめに

オペレーションズ・リサーチが得意とする企業や自治体内のさまざまな問題と同様に、スポーツの世界においても科学的な手法を用いて成功した例がいくつかある。例えば、メジャーリーグのオークランド・アスレチックスのビリー・ビーン GM は、セイバーメトリクスと呼ばれる手法を用いて、低予算で強豪チームを作り上げている。

数理科学分野において野球の打撃評価に関する最初の研究論文は、1959 年の Lindsey [1] であるとされ、統計的な手法を用いて野球を分析している。本稿の主題となる「野球への動的計画アプローチ」は、1960 年の Howard の名著 [2] に早くもその原型が示される。Howard は 1 インニングの期待得点最大化を評価基準として、25 状態のマルコフ決定過程 (Markov Decision Process) として定式化している。アウトカウントが無死・一死・二死、各塁に走者がいる・いないの組合せに 3 アウトチェンジを加えると、 $3 \times 2^3 + 1 = 25$ 状態となる。また、監督の指示として、打撃以外に盗塁と犠打が考慮されている。チームの打者 9 人の能力はすべて等しいと仮定し、推移確率 (犠打の成功率など) を人工的に設定したうえで、監督がとるべき最適政策を計算機を用いて実際に解いている。その後、Bellman [3] はより詳細な定式化を行い、打者ごとの解析のみならず、1 球ごとの解析についても提案している。Bellman は 1 インニングの期待得点最大化と、少なくとも k 点

を得点する閾値確率最大化の二つの基準について議論し、戦略に関する鋭い洞察を与えている。しかしながら、当時の計算機能力の不足もあり、現実のデータに適用されることはなかった。

一方、行列解析的なマルコフ連鎖アプローチを用いる二つの論文が 1977 年に発表される。D'Esopo and Lefkowitz [4] では、同一選手が繰り返し打席に立つと仮定するとき 1 インニングの期待得点を評価指標とするスコアリング・インデックス (SI) を提案している。Cover and Keilers [5] では、同様の指標として、同一選手が繰り返し打席に立つと仮定するとき 1 試合の期待得点を評価指標とする OERA (Offensive Earned-Run Average) 値を提案している。OERA モデルでは、野球のルールをある程度単純化して、吸収的マルコフ連鎖モデルを適用することにより、期待得点を計算する。Howard や Bellman が戦略の最適化に焦点を当てたのに対し、行列解析的なアプローチは個々の選手の貢献度を数値化することを目的としている。以降、この手法が人気を集めることになる。その後、1997 年の Bukiet et al. [6] では、打順を考慮し、9 人の選手がチームとして所与の順番で打席に立つと仮定するとき、1 試合の期待得点を計算するアルゴリズムが提案されている。また、勝率についても、一定の近似を用いた考察がなされている。

日本国内においても、1979 年の鳩山 [7] に始まり、さまざまな研究がなされている。Ano [8] や武井ら [9] は、単純化されている OERA モデルの野球ルールに、盗塁などのルールを加え、より現実的な推移率による期待得点を求めている。また、これらの研究では、日本国内の興味として、日本プロ野球のデータを用いて計算していることも、功績の一つであるといえる。さらに、廣津・宮地 [10] は、Bukiet et al. [6] のモデル

きら あきふみ

東北大学大学院経済学研究科

〒 980-8576 宮城県仙台市青葉区川内 27-1

いなかわ けいすけ

秋田県立大学システム科学技術学部

〒 015-0055 秋田県由利本荘市土谷字海老ノ口 84-4

を基礎として、SI と OERA に打順（ラインナップ）を考慮した SIL と OERAL を提案し、日本代表チームの最適打順について計算している。2008 年の大澤・合田 [11] は、1 試合の勝率について詳細に計算している。これより前の 1 試合の期待得点のほとんどは、1 イニングの期待得点を 9 倍、あるいはある程度の近似により計算しているものが主流であった。しかしながら、大澤・合田では、各イニングが n 番打者で終了する確率を計算することにより、詳細な期待得点と、延長戦も含めた勝率を計算している。

Howard の提案から半世紀、計算機的能力も飛躍的に発達し、再び動的計画アプローチに立ち返るときが来る。2008 年、Turocy [12] は、対戦相手チームの戦略も加味し、マルコフ決定過程のマルチ・エージェント拡張であるマルコフゲーム¹を用いてモデル化し、数値実験も行っている。対戦する両チームはそれぞれ、自チームが勝つ確率を最大化するという評価基準が用いられ、守備側の監督の指示として、敬遠も考慮されている。状態数は、約 213 万状態である。動的計画法の Bellman 方程式やアルゴリズムの詳細は省略されているが、1 試合のゲームの値、すなわち、両チームがベストを尽くしたときの均衡勝率を 1 分以内で高精度に解くことができると述べている。その後、Kira and Inakawa [15] は延長は最大で 12 回までとする日本プロ野球ルールを採用し、野球を約 350 万状態の有限マルコフゲームとして定式化し、ゲームの値とマルコフ完全均衡点を約 1 秒で計算している。ただし、敬遠を考慮していない。

本稿では、敬遠も考慮し、約 645 万状態の有限マルコフゲームとして定式化する。Turocy の定式化との微妙な違いと、それにより生じる重要な差異についても触れる。

2. マルコフゲームとしての定式化

野球の試合におけるさまざまな状況の一つ一つをここでは状態 (state) と呼ぶ。図 1 は一死一塁という状態から次の状態へ、またさらにその次の状態への推移の過程を表したものである。野球の 1 試合全体は巨大なゲームの木で記述することができるので、野球は展開形ゲーム (game in extensive form) と呼ばれるゲーム理論のクラスにも含まれる。

図 1 において、■で表された状態は守備側が意思決定を行う手番 (move) であり、打者と勝負するか敬遠

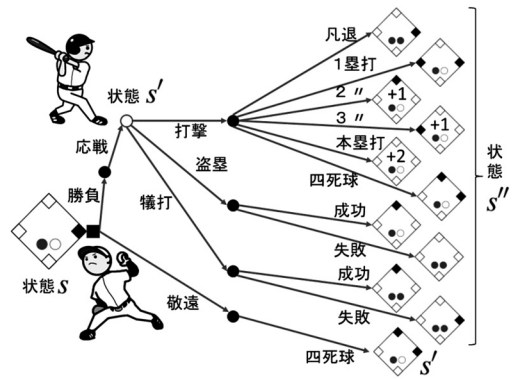


図 1 野球ゲームの木 (2 段階の状態推移のみを抜粋)

するかを決定する。○で表された状態は攻撃側が意思決定を行う手番であり、打撃、盗塁、犠打の中から指示を決定する。また、●は偶然手番 (chance move) と呼ばれるものであり、そこから出る枝の重み (推移確率) に従って、次の状態が確率的に決まる。ただし、出る枝が 1 本しかない偶然手番では、次の状態が確定的に定まる。また、偶然手番における推移確率は現在の状態とそこでの決定のみに依存し、それ以前の状態や決定の履歴には依存しないというマルコフ性 (Markov property) を仮定する。これは本モデルにおいて最も重要な仮定である。現実の野球では、例えば、相手守備陣に警戒されるので、三打席連続でセーフティバントを成功させることは難しく、成功率は過去の履歴に依存すると言ってよいだろう。したがって、マルコフ性の仮定は現実との誤差がある。しかしながら、計算の効率性とのトレードオフを考えると、マルコフ性を仮定することは十分に妥当と思われる。

さて、本稿では野球をマルコフゲームとして定式化する。マルコフゲームは、状態空間、決定空間、推移確率、利得関数といった構成要素でゲームを記述する。まず、野球における状態を定義する。便宜上、先攻チームと後攻チームをそれぞれチーム 0 とチーム 1 と呼ぶことにする。集合 S を野球におけるさまざまな状態を集めた状態空間 (state space) とし、各状態 $s \in S$ を七つの成分 $s = (\iota, \tau, \omega, \lambda, \mathbf{r}, \mathbf{b}, \mathbf{m})$ で表現する。各成分は次のように定義される。

1. $\iota \in \{1, 2, \dots, 12\}$ は現在のイニングを表す。 $\iota = 9$ が最終回であるが、同点の場合、最大で $\iota = 12$ まで延長戦が行われる。
2. $\tau \in \{0, 1\}$ は各回の表 (0)・裏 (1) を表す。
3. $\omega \in \{0, 1, 2, 3\}$ はアウトカウントを表す。
4. λ は得点差を表し、チーム 0 の得点からチーム 1

¹ Shapley [13] による提唱、Zachrisson [14] も有名。

の得点を引いた値を表す。

5. $\mathbf{r} = (r_3, r_2, r_1)$: $r_1 \in \{0, 1, \dots, 9\}$ は 1 塁走者がいなければ 0, いればその走者の打順の値をとる。
 $r_i \in \{0, 1\}$ ($i = 2, 3$) は i 塁走者がいなければ 0, いれば 1 の値をとる。
6. $\mathbf{b} = (b_0, b_1)$: $b_i \in \{1, 2, \dots, 9\}$ ($i = 0, 1$) はチーム i の打順がどの打者に回っているかを表す。攻撃時は現在の打者の打順を表し、守備時は次の攻撃の先頭打者の打順を表す。
7. $m \in \{0, 1\}$: この状態が先攻チームの手番 (0) か後攻チームの手番 (1) かを表す。

先攻チームと後攻チームの得点差を考慮する理由は、各状態における最適意思決定が得点差に依存するためである。例えば、9 回裏・無死一塁の場面で犠打をすべきか否かを考えてみよう。1 点取れば勝つ場合と 5 点取らなければ負ける場合とでは、犠打の有効性は違ってくるだろう。また、 \mathbf{r} の定義において、 r_1 のみを区別するのは、一塁から二塁への盗塁 (二盗) を行う際に、走者に依存した成功確率を考慮するためである。本稿では二塁から三塁への盗塁 (三盗) および三塁からホームベースへの盗塁 (ホームスチール) は考慮しないものとする。ここで、 $\mathcal{S}_Q \subset \mathcal{S}$ を試合終了の状態 (吸収状態) の全体とする。

$$\begin{aligned} \mathcal{S}_Q &:= \mathcal{S}_Q^1 \cup \mathcal{S}_Q^2 \cup \dots \cup \mathcal{S}_Q^5, \\ \mathcal{S}_Q^1 &:= \{s \in \mathcal{S} \mid \iota \geq 9, \tau = 1, \omega = 3, \lambda > 0\}, \\ \mathcal{S}_Q^2 &:= \{s \in \mathcal{S} \mid \iota = 9, \tau = 0, \omega = 3, \lambda < 0\}, \\ \mathcal{S}_Q^3 &:= \{s \in \mathcal{S} \mid \iota \geq 9, \tau = 1, \lambda < 0\}, \\ \mathcal{S}_Q^4 &:= \{s \in \mathcal{S} \mid \iota = 12, \tau = 1, \omega = 3, \lambda = 0\}, \\ \mathcal{S}_Q^5 &:= \{s \in \mathcal{S} \mid \lambda \leq -30\} \cup \{s \in \mathcal{S} \mid \lambda \geq 30\}. \end{aligned}$$

\mathcal{S}_Q^1 は先攻チームが勝利するケースである。 \mathcal{S}_Q^2 は 9 回表が終了した時点で、後攻チームの勝利が確定するケースである。 \mathcal{S}_Q^3 は後攻チームのサヨナラ勝ちに対応している。 \mathcal{S}_Q^4 は延長 12 回を終了し、引き分けとなるケースである。 \mathcal{S}_Q^5 はインングの途中であってもコールドゲーム成立を採用する mercy-rule [12] である。このルールは事実上、吸収確率が限りなくゼロに近い吸収壁であり、mercy-rule による誤差はないと考えてよい。さらに、延長を 12 回までとする日本プロ野球ルールとの組合せにより、ゲームの長さが有限となる。 $\omega = 3$ である状態 $s \notin \mathcal{S}_Q$ は 3 アウトチェンジ後の状態と同一視する。以上の定義により、試合開始 (プレイボール) 時の初期状態 s_0 は次のように表される。

$$\begin{aligned} s_0 &= (\iota, \tau, \omega, \lambda, \mathbf{r}, \mathbf{b}, m)_0 \\ &= (1, 0, 0, 0, (0, 0, 0), (1, 1), 1). \end{aligned}$$

\mathcal{S}_i をチーム i の手番である状態の集合とする。

$$\mathcal{S}_i := \{s = (\iota, \tau, \omega, \lambda, \mathbf{r}, \mathbf{b}, m) \in \mathcal{S} \setminus \mathcal{S}_Q \mid m = i\}.$$

$\mathcal{S}_0 \cup \mathcal{S}_1$ の中には、初期状態 s_0 から到達不可能な成分の組合せも含まれる。例えば、先攻チームの攻撃中に、打者が 4 番打者 ($b_0 = 4$) であるのに、一塁走者が 5 番打者 ($r_1 = 5$) という組合せは起こりえない。本モデルにおいて、到達可能な状態数を計算プログラムでカウントしたところ 6, 454, 296 状態であった。

本稿では、決定空間 (action space) として $\mathcal{A} := \{\text{勝負, 敬遠, 打撃, 盗塁, 犠打}\}$ を考える。各状態 $s \in \mathcal{S}_0 \cup \mathcal{S}_1$ で選択可能な決定全体を表す集合 $\mathcal{A}(s) \subset \mathcal{A}$ を可能決定空間 (feasible action space) と呼び、任意の状態 $s = (\iota, \tau, \omega, \lambda, \mathbf{r}, \mathbf{b}, m) \in \mathcal{S}_0 \cup \mathcal{S}_1$ に対して、次を満たすように定義する。

$$\begin{aligned} \text{勝負, 敬遠} \in \mathcal{A}(s) &\iff \tau \neq m, \\ \text{打撃} \in \mathcal{A}(s) &\iff \tau = m, \\ \text{盗塁} \in \mathcal{A}(s) &\iff \tau = m \wedge r_2 = 0 \wedge r_1 \geq 1, \\ \text{犠打} \in \mathcal{A}(s) &\iff \tau = m \wedge \omega \leq 1 \wedge \sum_i r_i \geq 1. \end{aligned}$$

ただし、“ \wedge ” は論理積 (かつ) を表す。すなわち、守備側にとって、打者と勝負するか敬遠するかは常に選択可能であり、攻撃側にとって打撃は常に選択可能である。また、盗塁は二塁走者がなく、一塁走者がいるときのみ選択可能である。犠打は走者が少なくとも 1 人いるときのみ選択可能である。三塁走者がいるときの犠打はスクイズを意味する。

決定 $a \in \mathcal{A}(s)$ を選択するとき、確率的に生じる結果の全体を $\mathcal{X}(a)$ とし、次のように定める。

$$\mathcal{X}(a) := \begin{cases} \{\text{応戦}\} & a = \text{勝負}, \\ \{\text{四球}\} & a = \text{敬遠}, \\ \{\text{凡打, 単打,} \\ \quad \text{二塁打, 三塁打,} & a = \text{打撃}, \\ \quad \text{本塁打, 四死球}\} \\ \{\text{成功, 失敗}\} & a = \text{盗塁, 犠打}. \end{cases}$$

$\mathcal{X}(a)$ は図 1 の偶然手番から出る枝のリストに相当する。 $p(x \mid s, a) \in [0, 1]$ は状態 s で決定 $a \in \mathcal{A}(s)$ が選択されるとき、結果 $x \in \mathcal{X}(a)$ が発生する確率を表す。

$$\sum_{x \in \mathcal{X}(a)} p(x|s, a) = 1, \quad \forall (s, a) \text{ s.t. } a \in \mathcal{A}(s).$$

$\mathcal{X}(a)$ 上の確率分布は、得点圏打率なども詳細に考慮することができる。本稿では単純化のため、確率分布は打撃・犠打・盗塁を行う選手のみ依存し、状態を構成するその他の成分には依存しないと仮定する。表 1 は日本プロ野球 2013 年シーズンの実績値をもとに作成した東北楽天イーグルス選手の確率パラメータである。ただし、企盗塁数と企犠打数が極端に少ない選手の盗塁成功率と犠打成功率は一部調整を行っている。具体的には、年間の企盗塁数が 4 に満たない選手の盗塁成功率は実際の値と .700 の小さい値を採用する。同様に、年間の企犠打数が 4 に満たない選手の犠打成功率は実際の値と .800 の小さい値を採用する。また、企犠打数が 0 の選手の成功率は .000 とする。

先行研究と同様に本稿でも野球を簡約化する。文献により微妙な差異があるが本稿では以下を用いる。

【簡約ルール】

1. 凡打によって走者も打者も進塁しない。
2. 単打は一塁走者を三塁へ進塁させ、二塁走者と三塁走者をホームへ生還させる。
3. 二／三塁打は全走者をホームへ生還させる。
4. ダブルプレーはないとする。
5. 盗塁成功時は一塁走者を二塁に進塁させる。
6. 盗塁失敗時は一塁走者をアウトにする。
7. 犠打が成功した場合は走者を一つ先の塁に進塁させ、犠打を行った打者はアウトとする。
8. 犠打が失敗した場合は最も本塁に近い走者がアウトになり、それ以外の走者を一つ先の塁に進塁させ、打者を一塁走者とする。

状態 s で決定 $a \in \mathcal{A}(s)$ を選択した結果が $x \in \mathcal{X}(a)$

表 1 東北楽天イーグルス選手の確率パラメータ

選手	打率	打撃						盗塁 成功	犠打 成功
		凡打	塁打				四球 ・死球		
			一	二	三	本			
1 岡島豪郎	.313	.606	.235	.034	.004	.004	.117	.500	1.000
2 藤田一也	.257	.697	.203	.032	.004	.002	.062	.600	.971
3 銀次	.317	.633	.234	.046	.006	.008	.073	.500	.333
4 ジョーンズ	.243	.605	.114	.035	.002	.044	.201	.571	*.000
5 マギー	.292	.621	.157	.051	.000	.048	.123	.400	*.000
6 栞田慎太郎	.271	.643	.157	.056	.000	.026	.118	*.700	*.800
7 松井稼頭央	.246	.688	.146	.053	.004	.022	.087	.500	*.800
8 嶋基宏	.248	.690	.192	.026	.002	.008	.081	.500	.941
9 聖澤諒	.278	.650	.204	.035	.008	.004	.100	.724	1.000

(出典) データで楽しむプロ野球 [16] より作成

であるとき、簡約ルールの下で次状態 s' は一意に定まる。すなわち、次状態は現在の状態 s 、そこでの決定 a 、結果 x の関数である。この関数を t とする。

$$s' = t(s, a, x).$$

任意の吸収状態 $s = (l, \tau, \omega, \lambda, \mathbf{r}, \mathbf{b}, m) \in \mathcal{S}_Q$ に対して、チーム i の利得 (payoff) $\psi_i(s)$ を次で定義する。

$$\psi_0(s) := \begin{cases} 1 & \lambda > 0, \\ 0 & \lambda \leq 0, \end{cases} \quad \psi_1(s) := \begin{cases} 0 & \lambda \geq 0, \\ 1 & \lambda < 0. \end{cases}$$

すなわち、試合が終了した時点で、勝ったチームのみが利得 1 を得ることができ、それ以外は利得 0 となる。

定義 2.1 (マルコフ政策). 写像 $\pi_i : \mathcal{S}_i \rightarrow \mathcal{A}$ が任意の $s \in \mathcal{S}_i$ に対して、 $\pi_i(s) \in \mathcal{A}(s)$ を満たすとき、 π_i をチーム i の(確定的)マルコフ政策 (Markov policy) という。チーム i のマルコフ政策全体を Π_i とする ($i = 0, 1$).

最後に、両チームの監督の目的関数はそれぞれ「自チームが勝つ確率」であるが、これを改めて定義しておこう。両チームがそれぞれマルコフ政策の中から $\pi_0 \in \Pi_0, \pi_1 \in \Pi_1$ を選び採用したとする。このとき、図 1 における■および○で表された手番での選択は採用したマルコフ政策により確定する。したがって、ゲームにおける状態変化は●で表された偶然手番の確率的推移のみに左右される。ゆえに、ゲーム開始時点の状態 s_0 から n 回推移したのちの状態を X_n とすると $\{X_n\}$ はマルコフ連鎖 (Markov chain) とみなすことができる。そこで、 $\{X_n\}$ の \mathcal{S}_Q への初到達時刻を T とする。

$$T := \min\{n \mid X_n \in \mathcal{S}_Q\} < \infty.$$

すなわち、試合終了までに偶然手番から出る枝を T 本通過したことを意味する。このとき、チーム i が勝つ確率 $v_i(s; \pi_0, \pi_1)$ が定まる。

$$v_i(s; \pi_0, \pi_1) := E^{\pi_0, \pi_1} [\psi_i(X_T) \mid X_0 = s], \quad i = 0, 1.$$

ただし、 E^{π_0, π_1} は両チームがそれぞれマルコフ政策 π_0, π_1 を採用した下での条件付き期待値を表す。このように、目的関数が確率であるときに、それを期待値の形に変換する方法は、マルコフ決定過程の分野でしばしば用いられる (例えば Kira et al. [17]).

3. マルコフ完全均衡と動的計画法

マルコフ完全均衡点とは任意の状態から始まる部分ゲームに対してナッシュ均衡 (Nash equilibrium) を導くマルコフ政策の組のことである。

定義 3.1 (マルコフ完全均衡点). (確定的) マルコフ政策の組 $(\pi_0^*, \pi_1^*) \in \Pi_0 \times \Pi_1$ が (純戦略) マルコフ完全均衡点 (Markov perfect equilibrium) であるとは, (π_0^*, π_1^*) が部分ゲーム完全均衡点であることである。

$$v_0(s; \pi_0, \pi_1^*) \leq v_0(s; \pi_0^*, \pi_1^*), \quad \forall s \in \mathcal{S}, \quad \forall \pi_0 \in \Pi_0, \\ v_1(s; \pi_0^*, \pi_1) \leq v_1(s; \pi_0^*, \pi_1^*), \quad \forall s \in \mathcal{S}, \quad \forall \pi_1 \in \Pi_1.$$

定義 3.2 (ゲームの値). (π_0^*, π_1^*) をマルコフ完全均衡点とする. このとき, 両チームの均衡勝率

$$(V_0(s), V_1(s)) := (v_0(s; \pi_0^*, \pi_1^*), v_1(s; \pi_0^*, \pi_1^*))$$

を状態 s におけるゲームの値と呼ぶ。

正確には完全情報という概念を持ち出す必要があるが, 同時手番がない有限マルコフゲームに対しては, 純戦略マルコフ完全均衡点が少なくとも一つ存在することがよく知られている. 理由はこれから述べる後ろ向きの帰納法 (backward induction) を用いて純戦略マルコフ完全均衡点を構成できるからである. さらに, マルコフ完全均衡点が複数存在する場合でも, ゲームの値は均衡点に依存しない. 一方, Turocy [12] は延長戦が無制限の MLB ルールを用いているので, 状態に依存したある確率で盗塁をするといった混合戦略マルコフ完全均衡点の存在定理を利用している. 実際, Turocy のモデルも状態数・決定数が有限であるため, 確定的マルコフ政策の総数も有限である. ゆえに, 戦略形ゲーム (game in strategic form) に帰着でき, 混合戦略のクラスで決着する. したがって, 「日本とメジャーの違いはなんですか?」という野球記者の質問に対する OR 的回答の一つは「9 回以降の作戦に関する純戦略マルコフ完全均衡点の有無」である。

さて, ゲームの値とマルコフ完全均衡点が満たす再帰式を導出する. 帰納法の初期条件として, 試合終了の状態 $s_Q \in \mathcal{S}_Q$ におけるゲームの値は明らかに, $V_i(s_Q) = \psi_i(s_Q)$ である. そこで, 図 2 の○の手番における攻撃側の意思決定について考えよう. ○の手番から推移可能なすべての状態 s'' について, ゲームの値 $V_i(s'')$ を計算済みと仮定する. このとき, 各 $V_i(s'')$

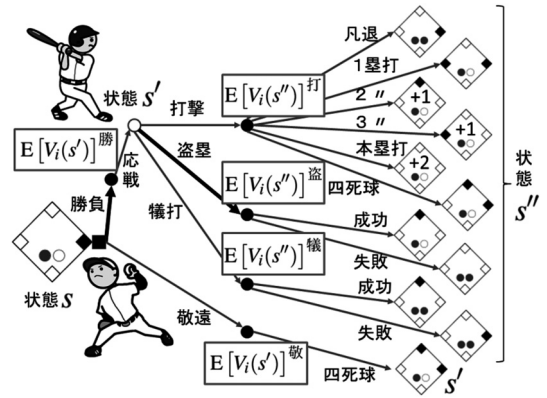


図 2 後ろ向きの帰納法

に偶然手番から出る枝の重みを掛けて足し合わせる, すなわち加重平均をとることで, 攻撃側が打撃, 盗塁, 犠打を選択したときのチーム i の勝率をそれぞれ計算できる. 攻撃側は自チームの勝率が最大となる決定を選ぶはずである. 次に, ■の手番における守備側の意思決定について考えよう. ■の手番から推移可能なすべての状態 s' について, ゲームの値 $V_i(s')$ を計算済みと仮定する. 同様に加重平均をとることで, 守備側が勝負, 敬遠を選択したときのチーム i の勝率をそれぞれ計算できる. 守備側は自チームの勝率が最大となる決定を選ぶはずである. 以上の考察により, 次の定理を得る。

定理 3.1 (Bellman 方程式). ゲームの値関数 $V_i : \mathcal{S} \rightarrow [0, 1]$ およびマルコフ完全均衡点 $(\pi_0^*, \pi_1^*) \in \Pi_0 \times \Pi_1$ は次の再帰式を満たす.

$$V_i(s) = \begin{cases} \psi_i(s) & s \in \mathcal{S}_Q, \\ \max_{a \in \mathcal{A}(s)} \sum_{x \in \mathcal{X}(a)} V_i(t(s, a, x)) p(x|s, a) & s \in \mathcal{S}_i, \\ \sum_{x \in \mathcal{X}(\pi_j^*(s))} V_i(t(s, \pi_j^*(s), x)) p(x|s, \pi_j^*(s)) & s \in \mathcal{S}_j, \end{cases}$$

$$\pi_i^*(s) \in \arg \max_{a \in \mathcal{A}(s)} \sum_{x \in \mathcal{X}(a)} V_i(t(s, a, x)) p(x|s, a), \quad s \in \mathcal{S}_i.$$

ただし, $(i, j) = (0, 1), (1, 0)$.

各状態から推移可能な次の状態は高々 10 通りであり, 推移後の勝率と推移確率との加重平均を定数時間で計算できる. したがって, 本稿で定式化した野球は, 状態数に対して, 線形時間で解くことができる。

4. 後攻チームが有利って本当!?

野球に関して, 「後攻が有利」などという会話をよ

表 2 ホーム・ビジター別の勝敗表 (2013 シーズン)

チーム	試合	ホーム				ビジター			
		勝	負	分	勝率	勝	負	分	勝率
巨人	144	46	21	5	.639	38	32	2	.528
阪神	144	36	33	3	.500	37	34	1	.514
広島	144	34	37	1	.472	35	35	2	.486
中日	144	29	42	1	.403	35	35	2	.486
DeNA	144	28	44	0	.389	36	35	1	.500
ヤクルト	144	33	38	1	.458	24	45	3	.333
楽天	144	42	28	2	.583	40	31	1	.556
西武	144	42	29	1	.583	32	37	3	.444
ロッテ	144	41	29	2	.569	33	39	0	.458
ソフトバンク	144	37	34	1	.514	36	35	1	.500
オリックス	144	32	38	2	.444	34	35	3	.472
日本ハム	144	35	36	1	.486	29	42	1	.403
計		435	409	20	.503	409	435	20	.473

(出典) プロ野球マルチデータ置き場 [18] より作成

(注) 勝率は全試合数に占める勝利数の割合

表 3 試合開始時におけるゲームの値と作戦の効果

後攻 先攻	---	一敬	一犠一	一犠敬	盗---	盗一敬	盗犠一	盗犠敬
---	.4932	.4931	.4843	.4843	.4889	.4889	.4781	.4780
一敬	.4932	.4932	.5037	.5037	.4980	.4980	.5107	.5107
一犠一	.4933	.4933	.4860	.4859	.4892	.4892	.4817	.4816
一犠敬	.4931	.4931	.5019	.5020	.4976	.4976	.5069	.5069
盗---	.5014	.5002	.4926	.4914	.4971	.4960	.4864	.4852
盗一敬	.4862	.4872	.4966	.4975	.4910	.4919	.5035	.5045
盗犠一	.5016	.5004	.4942	.4930	.4975	.4963	.4899	.4887
盗犠敬	.4861	.4870	.4949	.4958	.4905	.4915	.4997	.5007
盗---	.4975	.4973	.4887	.4885	.4932	.4931	.4824	.4823
盗一敬	.4893	.4894	.4997	.4998	.4941	.4942	.5068	.5068
盗犠一	.4976	.4975	.4903	.4901	.4936	.4934	.4860	.4858
盗犠敬	.4892	.4892	.4980	.4981	.4936	.4937	.5029	.5030
盗---	.5072	.5046	.4984	.4957	.5029	.5003	.4923	.4895
盗一敬	.4812	.4832	.4915	.4936	.4859	.4880	.4983	.5005
盗犠一	.5074	.5047	.5000	.4973	.5033	.5006	.4957	.4930
盗犠敬	.4811	.4831	.4899	.4919	.4855	.4876	.4947	.4968

※上段は先攻, 下段は後攻チームの均衡勝率

く耳にする。日本プロ野球ではビジターチームが先攻, ホームチームが後攻として試合が行われる。2013 年シーズンのホーム・ビジター別勝敗は表 2 のようになっている。

セ・リーグとパ・リーグを合せて合計 864 試合が行われるが, 先攻チームの勝率は .473, 後攻チームの勝率は .503 であり, 後攻の勝率が 3%ほど上回っている。観客の声援などホームスタジアムで試合ができるアドバンテージは種々考えられるが, 野球のルールそのものに起因する優位性はあるのだろうか。

Turocy [12] は同一のチームが対戦する仮想的な試合に対してゲームの値を計算することで, 後攻の優位性を議論している。その際, 監督が指示できる作戦として, 盗塁・犠打・敬遠の ON, OFF を切り替え, 計 8 通りの状況下でゲームの値を比較している。本稿では, 表 1 の東北楽天イーグルス同士を対戦させ, 先攻チームと後攻チームのそれぞれに対して監督が指示できる作戦の ON, OFF を切り替え, 計 64 通りの状況下で数値実験を行った。その結果を表 3 に示す。

両チームの作戦をすべて ON にしたときの計算時間が最長であり, そのとき, 1.61 秒/1 試合で純戦略マルコフ完全均衡点とゲームの値の計算が完了した。実行環境は Intel® Core™ i7-3770K 16GB メモリ搭載のデスクトップ型 PC である。

両チームともに敬遠のみが ON の状況下では先攻チームの勝率が高い。逆に, 盗塁のみが ON もしくは犠打のみが ON の状況下では, 後攻チームの勝率が高い。したがって, 敬遠は先攻チームに, 盗塁および犠打は後攻チームにそれぞれ有利に働く。ゆえに, 両チームの作戦をすべて ON にしたときに, 先攻と後攻どちらのチームの勝率が高いかは偶然手番の推移確率に依存する。しかしながら, 表 3 のとおり, 敬遠の影響は盗塁と犠打よりも小さく, 通常は後攻チームの勝率が高くなると考えてよいだろう。以上の結果は Turocy が示した事実と一致する。ただし, Turocy は「違いは僅かであり大した差はない」とも述べている。確かにそのとおりであるが, .007 の勝率増が 144 試合のペナントレースでは 1 試合の勝利数増になることを考慮すれば, 見過ごせないかもしれない。

さて, 第 2 節で状態を定義する際に, 試合に勝つためには相手チームとの得点差を考慮しなければならないことを述べた。すなわち, 相手チームが何点得点したかという情報が大切になる。打撃・盗塁・犠打の意思決定を行う際に, 野球のルール上, 後攻チームのほうが常に 1 インニング分多く, 相手チームが何点得点したかを観測できる。逆に, 敬遠の意思決定に関しては, 先攻チームがある意味「後攻め」である。この非対称性こそが先攻と後攻の勝率に僅かな差をもたらす原因と考えられる。

5. 最適打順

Bukiet et al. [6] の論文以降, 行列解析的なマルコフ連鎖アプローチにおけるホットトピックは最適打順の計算であり, その後の多くの論文でも計算されている。動的計画アプローチでは, 戦略の最適化自体に計算時間がかかるため, 打順を総当たりし, 最適打順を求める計算コストは極めて高い。ただし, 実際には 8! 通りの総当たりでよい。計算結果をうまく再利用すれば, 打順 $\sigma = (1, 2, 3, 4, 5, 6, 7, 8, 9)$ によるゲームの値を計算する時間と, 回転によって得られる打順 (例えば, $\sigma' = (2, 3, 4, 5, 6, 7, 8, 9, 1)$) によるゲームの値を全 9 通り計算する時間の合計はほとんど変わらない。

表 4 は東北楽天イーグルスの最適打順を計算した結果である。ただし, ここでも同一チームによる仮想

表 4 最適打順

選手	打率	四球		盗塁	犠打	後攻の打順		
		死球	成功	成功	標準	最悪	最適	
岡島豪郎	.313	.117	.500	1.000	1	8	1	
藤田一也	.257	.062	.600	.971	2	1	7	
銀次	.317	.073	.500	.333	3	3	4	
ジョーンズ	.243	.201	.571	.000	4	9	2	
マギー	.292	.123	.400	.000	5	7	3	
栴田慎太郎	.271	.118	.700	.800	6	6	5	
松井稼頭央	.246	.087	.500	.800	7	5	9	
嶋基宏	.248	.081	.500	.941	8	2	8	
聖澤諒	.278	.100	.724	1.000	9	4	6	
先攻(標準) v.s. 後攻における後攻勝率					.4968	.4824	.5067	

的な対戦とし、先攻チームの打順は表 1 で与えられるデフォルトの打順に固定した。その上で、後攻チームの打順を総当たりし、後攻チームの勝率が最も高くなる最適打順と、最も低くなる最悪打順を求めた。また、ゲーム開始時点におけるゲームの値が高精度に計算できれば十分であるので、mercy-rule を 30 点差でコールドゲームから 20 点差でコールドゲームに引き下げて数値実験を行った。1 スレッドで単純な総当たりを実行したため、計算に半日を要した。

両チームが試合中の意思決定に関してベストを尽くすという前提においては、最適打順と最悪打順の勝率差は僅か 2.43% である。ただし、144 試合のペナントレースに換算すると、勝利数の差は 3.49 である。勝利数が増える分だけ他チームの敗北数が増えることを考慮すると、ゲーム差はそれ以上となるだろう。

6. おわりに

本稿で紹介したモデルはまだまだ発展途上である。「ダブルプレーの回避も犠打の目的の一つであるので、簡約ルールに加えるべきでは？」と早速、野球ファンとしての著者自身から指摘を入れられてしまった。幸いにも計算コストの面ではまだまだモデルを拡張する余裕がある。本誌をお読みいただいた皆様に多くの知見をご教示いただければ幸いである。

参考文献

- [1] Lindsey, G. R., “Statistical Data Useful for the Operation of a Baseball Team,” *Operations Reserach*, **7**, 197–207, 1959.
- [2] Howard, R. A., *Dynamic Programming and Markov Processes*, M.I.T. Technology Press and Wiley, Cambridge, Mass, 1960.
- [3] Bellman, R., “Dynamic Programming and Markovian Decision Processes, with Application to Baseball,” *Optimal Strategies in Sports*, S. P. Ladany and R. E. Macol (eds.), Elsevier-North Holland, New York, 77–85, 1977.
- [4] D’Esopo, D. A. and Lefkowitz, B., “The Distribution of Runs in the Game of Baseball,” *Optimal Strategies in Sports*, S. P. Ladany and R. E. Macol (eds.), Elsevier North-Holland, 55–62, 1977.
- [5] Cover, T. M. and Keilers, C. W., “An Offensive Earned-Run Average for Baseball,” *Operations Research*, **25**, 729–740, 1977.
- [6] Bukiet, B., Harold, E. R. and Palacios, J. L., “A Markov Chain Approach to Baseball,” *Operations Research*, **45**, 14–23, 1997.
- [7] 鳩山由紀夫, “野球の OR,” *オペレーションズ・リサーチ*, **24**, 203–212, 1979.
- [8] Ano, K., “Modified offensive earned-run average with steal effect for baseball,” *Applied Mathematics and Computation*, **120**, 279–288, 2001.
- [9] 武井貴裕, 瀬古進, 穴太克則, “野球の最適打順を考えてみよう,” *オペレーションズ・リサーチ*, **47**, 142–147, 2002.
- [10] 廣津信義, 宮地力, “野球チームのラインナップ選定のための数理的な手法—日本代表チームの選定を例として—,” *オペレーションズ・リサーチ*, **49**, 380–389, 2004.
- [11] 大澤清, 合田憲人, “野球における走者の進塁状況を考慮した勝率計算方法,” *日本応用数学会論文誌*, **18**, 321–346, 2008.
- [12] Turocy, T. L., “In Search of the “Last-Ups” Advantage in Baseball: A Game-Theoretic Approach,” *Journal of Quantitative Analysis in Sports*, **4**(2) Article 5, 2008.
- [13] Shapley, L. S., “Stochastic games,” *Proceedings of the National Academy of Sciences of the United States of America*, **39**, 1095–1100, 1953.
- [14] Zachrisson, L. E., “Markov games,” *Annals Mathematics Studies: Advances in Game Theory*, M. Drescher, L. S. Shapley and A. W. Tucker (eds.), Princeton University Press, Princeton, **52**, 211–253, 1964.
- [15] Kira, K. and Inakawa, K., “On Markov perfect equilibria in baseball,” *Bulletin of Informatics and Cybernetics*, to appear.
- [16] データで楽しむプロ野球, <http://baseballdata.jp/> (Accessed 2014 Apr. 7).
- [17] Kira, K., Ueno, T. and Fujita, T., “Threshold probability of non-terminal type in finite horizon Markov decision processes,” *Journal of Mathematical Analysis and Applications*, **386**, 461–472, 2012.
- [18] プロ野球ヌルデータ置き場, <http://lcom.sakura.ne.jp/NulData/> (Accessed 2014 Apr. 7).