

数学とデータを使ってスポーツを科学する

—サッカー選手の価値評価とチーム戦術評価—

東京理科大学 創域理工学部 経営システム工学科 准教授 ことう まこと 後藤 允

■ スポーツがもつ可能性

スポーツとは、何でしょうか？ ある人にとっては、競技そのもの、つまり自分がするものでしょう。不断の努力によって実力を向上させ、よい成績をおさめることが目標になります。またある人にとっては、観戦して楽しむものかもしれません。応援しているチームや選手から夢と感動をもらい、明日への活力となる、これもスポーツの素晴らしい側面です。

もう少しライトに考えてみると、ジョギングをしたり、ジムで汗を流したり、週末にテニスをしたり、これらもちろんスポーツです。競技というよりは健康のためのスポーツで、生活を幸福で豊かにしてくれるでしょう。また、小学校や中学校、町内の運動会などもスポーツといえます。これらは友達との交流や、地域の一体感の醸成に資するものでしょう。

また視点を変えてみましょう。スポーツをするためには、道具やウェアが必要となりますし、場所も必要になってきます。つまり、環境を整備することが不可欠で、それには資金が必要となります。プロスポーツとなれば、必要な資金や消費される金額も桁違いに大きくなります。このように、スポーツは経済としての側面も重要になります。

以上のように、スポーツは「競技」「健康」「交流」「経済」という多くの側面をもち、多様な立場の人々にとって重要なものであるといえます。スポーツ基本法にも「スポーツとは、世界共通の人類の文化である」と書かれているように、スポーツは単に運動という言葉では収まらない、大きな可能性を秘めていることが分かります。本稿では、このように大きな可能性を秘めたスポーツに対して、数学とデータを使って科学的に分析している後藤研究室の取り組みを紹介します。

■ スポーツとデータ

まずは、データを使った分析から始めましょう。スポーツのデータ分析といえば、野球のセイバーメトリ

クスがそのはしりであるといわれています。映画「マネーボール」でも取り上げられたことで、一躍有名になった手法です。

そもそも野球は、データとの相性が非常によいスポーツです。攻撃と守備が明確に分かれていること、全てのプレイは投手が打者に対してボールを投げることから始まること、打者は順番に攻撃機会が回ってくることなど、場面の特定が容易であることがその理由です。したがって、打率、本塁打、打点、盗塁、防御率、守備率、奪三振など、さまざまな統計データを容易に収集することができます。

攻撃や守備の場面が順番に決まっている野球を「静的な」スポーツとすれば、選手が広いピッチを走り回り、攻守が目まぐるしく入れ替わるサッカーは「動的な」スポーツといえます。データが取りやすい静的なスポーツに比べて、動的なスポーツはデータが取りにくく、サッカーは野球ほどデータ分析が発展してこなかった歴史があります。しかし、近年のIT技術の向上やデータサイエンスの発展のおかげで、サッカーのデータ収集も急速に進歩しています。

具体的には、GPSを使った選手の位置情報取得によって、プレイ位置はもちろん、移動速度や走行距離を収集できるようになりました。また、心拍数や血中酸素濃度などのバイタルデータを取得することで、身体への負荷も数値化できるようになりました。ただし、これらのデータ取得にはウェアラブル端末の着用が必要で、選手によっては嫌がる場合がある他、取得データは基本的にチーム外には公開されないという欠点があります。

これらの問題点を解決するのが、映像解析技術です。試合の映像データから【表1】のようなプレイデータを抽出し、分析を可能にします。これらのデータはData Stadium社やStatsBomb社から提供されていますが、現状ではマンパワーによる手動抽出がなされているようです。AIによる自動抽出も研究が進められているようですが、実用にはまだ時間がかかるようです。

■ プロセスクラスタリングによる 攻撃戦術の評価

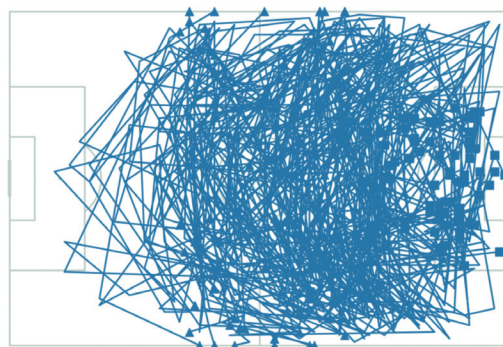
それでは、【表 1】のプレイデータを使った、サッカーチームの攻撃戦術分析を紹介します。まず、ボール保持からボールロストまでのポゼッションを 1 つのプロセスとして捉えます。【表 1】のオレンジ部分がチーム 1 の 1 つのプロセスになります。プレイヤー 1A からポゼッションが始まり、プレイヤー 1C がシュートを撃って外れたことでポゼッションが終わっています。続いて、チーム 2 のポゼッションがキーパー 2A のゴールキックから始まることになります。

このプロセスをピッチ上の座標にプロットすれば、ボールがどのように動いてどこでポゼッションが終わったかが視覚的に分かることになります。特に、有効な攻撃戦術を分析するために、シュートで終わったプロセスだけに限定します。あるチームの 1 シーズンにおけるシュートで終わった全プロセスをプロットしたのが【図 1】です。ピッチ上にぐちゃぐちゃに線が書かれていて、これでは何が何だか理解できません。

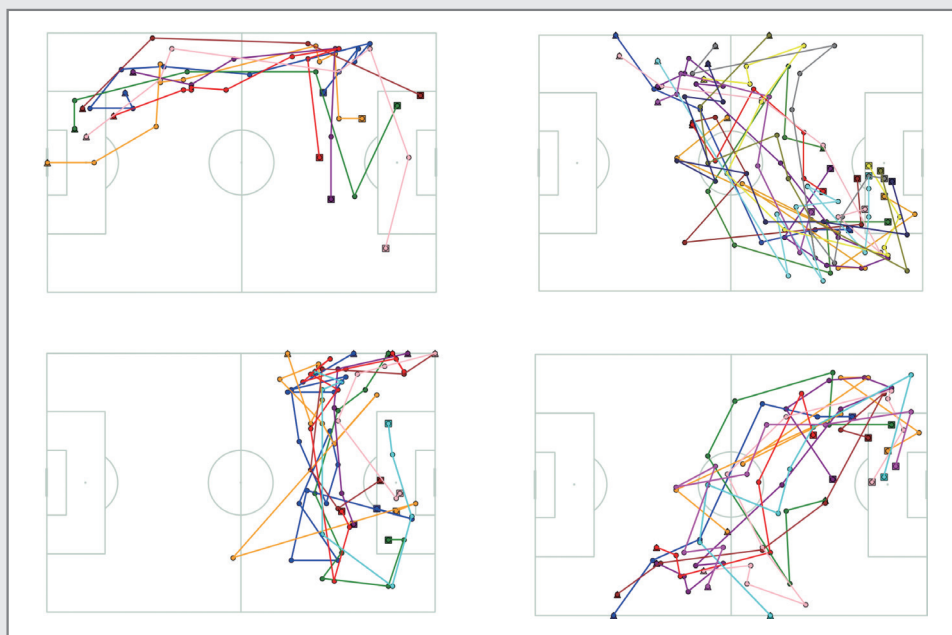
そこで、ボールの移動経路が類似したプロセス毎にまとめることを考えます。このような手法をクラスタリングといい、ここでは動的時間伸縮法 (Dynamic Time Warping; DTW) という手法を使います。DTW によってクラスタリングされたプロセスの例が、【図 2】です。【図 1】と比べて、類似したプロセスに分けられていることが分かります。

time	event	team	player	x	y
04:08.3	パス	1	1A	61.0	40.1
04:10.9	パス	1	1B	77.3	39.6
04:11.7	ドリブル	1	1C	83.0	23.6
04:13.2	シュート	1	1C	91.2	27.8
04:32.4	ゴールキック	2	2A	6.0	36.0

【表 1】サッカーのプレイデータ。パスやドリブルなどのイベントが発生した時刻、チームとプレイヤー、ピッチ上の座標が抽出されている。



【図 1】あるチームの 1 シーズンにおけるシュートで終わった全プロセス。三角がプロセス開始座標で、四角がプロセス終了座標。



【図 2】DTW によってクラスタリングされたプロセスの例。(左上) 左サイドを使った攻撃。(左下) アタッキングサードをワイドに使った攻撃。(右上) 左サイドから右サイドへ展開した攻撃。(右下) 右サイドから左サイドへ展開した攻撃。

戦術グループ	ケース数	ゴール数	xG	ゴール成功率	ケースごとの xG
start center and to center	29	6	3.70	0.207	0.128
start left and to left	169	26	19.81	0.154	0.117
start right and to right	185	20	19.84	0.108	0.107
start center and to right	50	4	5.23	0.080	0.105
start center and to left	51	9	5.06	0.176	0.099
start left and to right	21	1	2.07	0.048	0.099
start left and to center	78	7	7.41	0.090	0.095
start right and to left	30	5	2.59	0.167	0.086
start right and to center	42	0	0.63	0.000	0.015

【表 2】 あるチームの攻撃開始位置と方向によるプロセス分類と平均 xG.

ただし、クラスタリングされたすべてのプロセスが戦術的に解釈できるものとは限りません。速攻・遅攻、中央突破・サイド攻撃など、分析したい戦術を決めて、この戦術を規定するためのファクターを設定します。例えば、攻撃開始位置と攻撃方向をファクターとすれば、クラスタ毎に中央突破かサイド攻撃に分類できます。また、プロセスに要した時間から、速攻か遅攻に分類できます。このように、分析したい戦術にプロセスを分類することができます。

次に、攻撃戦術の評価方法を考えましょう。良い戦術とは何でしょうか？ 多くのゴールを決めていれば、それは良い戦術でしょう。しかし、優秀なストライカーがいれば、少々難しい場面でもゴールを決めてくれるでしょうし、逆に相手キーパーが優秀なら、完璧に崩れていても止められるかもしれません。そこで、シュートを撃った状況からゴールが決まる期待値を指標化した xG を使います。xG が 1 なら確実にゴールが決まるシュート、0 なら絶対に決まらないシュートということです。この指標なら選手の能力に関係なく、その戦術によって、どこまでゴールチャンスを作れたかを評価することができます。

あるチームの攻撃プロセスを攻撃開始位置と方向によって分類し、クラスタ毎に平均 xG を求めたのが【表 2】です。このチームの平均 xG が上位の戦術をみると、すべて開始位置から同じ方向への攻撃になっています。特に、右サイド攻撃と左サイド攻撃はケース数も多く、サイド攻撃を得意とするチームだということが明確に分かります。反対に、逆サイドへの展開は平均 xG が低く、苦手としていることが分かります。

このように、【表 1】のプレイデータをプロセスとして分析し、平均 xG を計算することで、チームが得意または苦手とする攻撃戦術を数値的に明らかにする

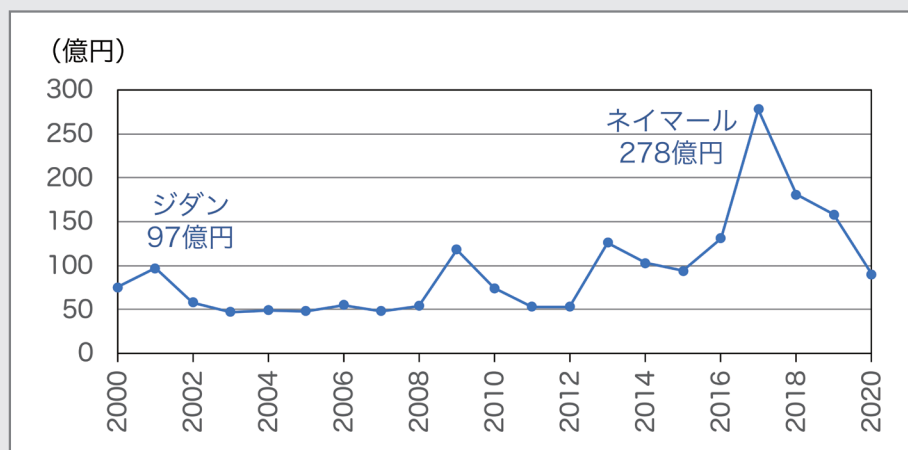
ことができます。この結果は、自チームの強化にも使えますし、対戦相手の戦術分析にも使えます。自チームの指向する戦術の平均 xG が高くなっていけば、攻撃がうまくいっているが、低ければ戦術練習の強度を上げる必要があります。また、対戦相手の平均 xG が高い攻撃戦術に対する守備戦術を準備すれば、相手のゴールチャンスを減らすことができます。

■ サッカー選手の金銭的価値

次に、サッカー選手の金銭的価値評価を紹介します。プロスポーツ選手の金銭的価値とは、何でしょうか？ MLB ロサンゼルスドジャースの大谷翔平選手は、年俸総額約 1,000 億円（為替レートは当時のもの）の 10 年契約を結びました。まさにとんでもない金額ですが、サッカーの世界では、移籍金というものが存在します。

選手は所属クラブと契約しているため、契約期間中は自由に移籍することができません。他クラブが選手を獲得するためには、移籍金を支払う必要があります。どうしても獲得したい選手には高い移籍金を支払う必要があるため、この移籍金がサッカー選手の金銭的価値を表していると考えられます。

【図 3】は、2000 年から 2020 年までの移籍金年間最高額の推移を表しています。2001 年にジダン選手がユベントスからレアルマドリーに移籍した際に 97 億円を記録し、その 16 年後にはバルセロナからパリサンジェルマンに移籍したネイマール選手が、約 3 倍の 278 億円を記録しています。近年のサッカー界では移籍金の高騰が顕著となっており、後藤研究室では、選手の適正な金銭的価値はどのように求めればよいのか、という問題に注目しています。



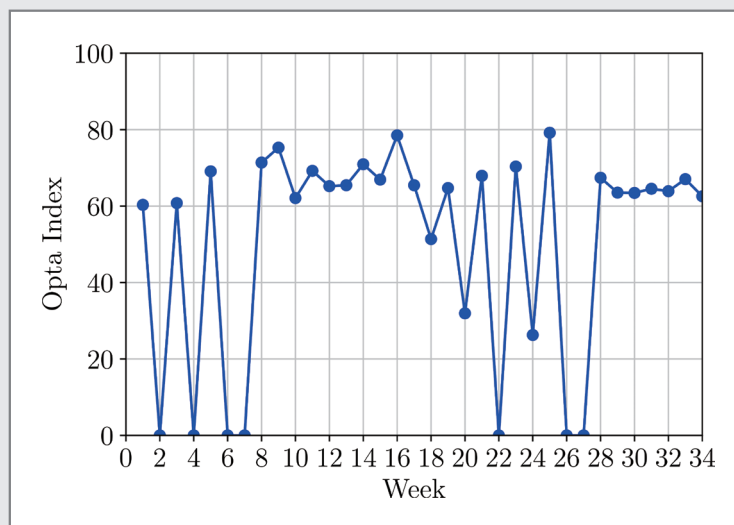
【図3】サッカー選手の移籍金年間最高額の推移 (為替レートは当時のもの)。

■ 金融工学的アプローチによる評価

2005年に発表された論文 Tunaru et al.³⁾では、選手の契約価値を金融工学的アプローチから計算するという、画期的な方法が提案されました。キーポイントは、選手のパフォーマンスとクラブの売上を結びつけることで、選手の金銭的価値を表すという点です。サッカー選手のパフォーマンスを表す指標として、Stats Perform社が提供するオプタ指数というものがあります。1試合で対象選手が積み重ねたプレイ毎に加点または減点し、客観的に点数化された指標です。クラブが試合で売り上げた金額を、選手毎にこのオプタ指数で按分すると、対象選手がその試合のパフォーマンスによって生み出した売上が計算されるということです。

J1リーグにおける、ある選手の1年間のオプタ指数の推移を表したのが【図4】です。時間経過とともに、不規則に上下していることが分かります。この時間変化は、金融工学における株価の変動によく似ているため、金融工学的アプローチが使えるということです。

金融商品にデリバティブというものがあり、多様な条項が付与された契約になっています。例えば、対象の株を1カ月後に1万円で売却する、または2カ月後に5千円で購入してもよい、などです。さらに、期間内に1万円を下回ってから権利が発生する、または7千円を超えたらその権利が消滅する、などの複雑な条項が付与されている場合もあります。これらの条項を数式でモデル化し、数学を駆使して契約の適正価格を求めるのが金融工学です。

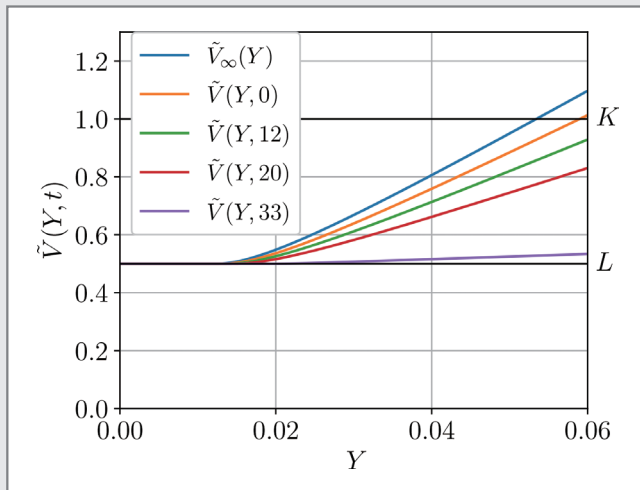


【図4】ある選手の1試合あたりのオプタ指数。

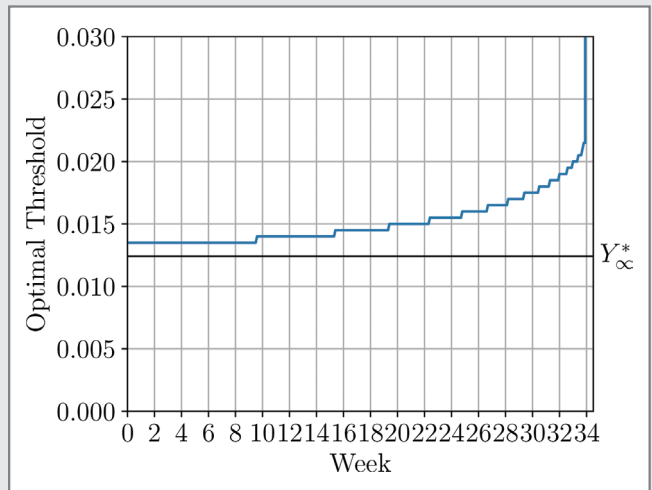
■ 移籍と放出を考慮した サッカー選手の価値評価

Tunaru et al.³⁾は2005年に発表された古い論文ですが、その後の発展が止まってしまっています。後藤研究室では、この研究を改良し、選手の移籍と放出を考慮したモデルを開発しました²⁾。選手の移籍に関しては、契約解除条項というものがあり、設定された移籍金を他クラブから提示された場合、クラブは選手を移籍させなければなりません。選手全員に契約解除条項が付与されているわけではありませんが、有名選手や人気選手には必ず付与されています。

また、選手のパフォーマンスが著しく低下した場合は、クラブはその選手を放出して移籍金を得ることもできます。この場合の移籍金は契約解除条項に設定されている金額よりも低くなりますが、レギュラークラスの選手ならば最低このくらいの水準という金額があります。さらに、サッカーには怪我がつきものであり、



【図5】残り契約期間毎の選手の契約価値。横軸 Y は選手がその試合のパフォーマンスによって生み出した売上、 K は契約解除条項の移籍金、 L は放出時の移籍金。系列は上から順に、契約が無期限とした場合、シーズン開始、残り 22 試合、14 試合、1 試合。



【図6】選手の放出閾値。各時点でのこの値を下回ったときに選手を放出する。

選手は怪我をすると試合を欠場することになるため、価値の損失が発生します。

これらを数式でまとめると、選手の契約価値は

$$V(Y, t) = \sup_{\theta \in \mathcal{T}} \mathbb{E} \left[\begin{aligned} & \int_t^{\eta \wedge \theta \wedge T} e^{-\rho(s-t)} Y_s ds \\ & - \sum_{\tau_i < \eta \wedge \theta \wedge T} e^{-\rho(\tau_i-t)} m Y_{\tau_i} \\ & + e^{-\rho(\eta-t)} K \mathbf{1}_{\{\eta < \theta \wedge T\}} \\ & + e^{-\rho(\theta-t)} L \mathbf{1}_{\{\theta \leq \eta \wedge T\}} \end{aligned} \right]$$

1. 毎試合発生する売上
2. 怪我による損失
3. 移籍による収益
4. 放出による収益

となり、金融工学的アプローチから計算した結果が【図5】です。ただし、使用データは仮想パラメータとなっています。

選手のパフォーマンスが高いほど契約価値が高く、時間が経過して契約終了に近づくほど契約価値が低くなっていくことが分かります。これは、契約が終了すると移籍金なしで選手が自由に移籍してしまうことを反映しています。また、選手のパフォーマンスが低下した場合は、クラブは選手を放出して移籍金を得ることができるため、契約価値は放出時移籍金を下回ることがないことも分かります。

【図6】は、選手の放出閾値を表しています。閾値とは、その値を境に状態が変わる値のことで、ここで

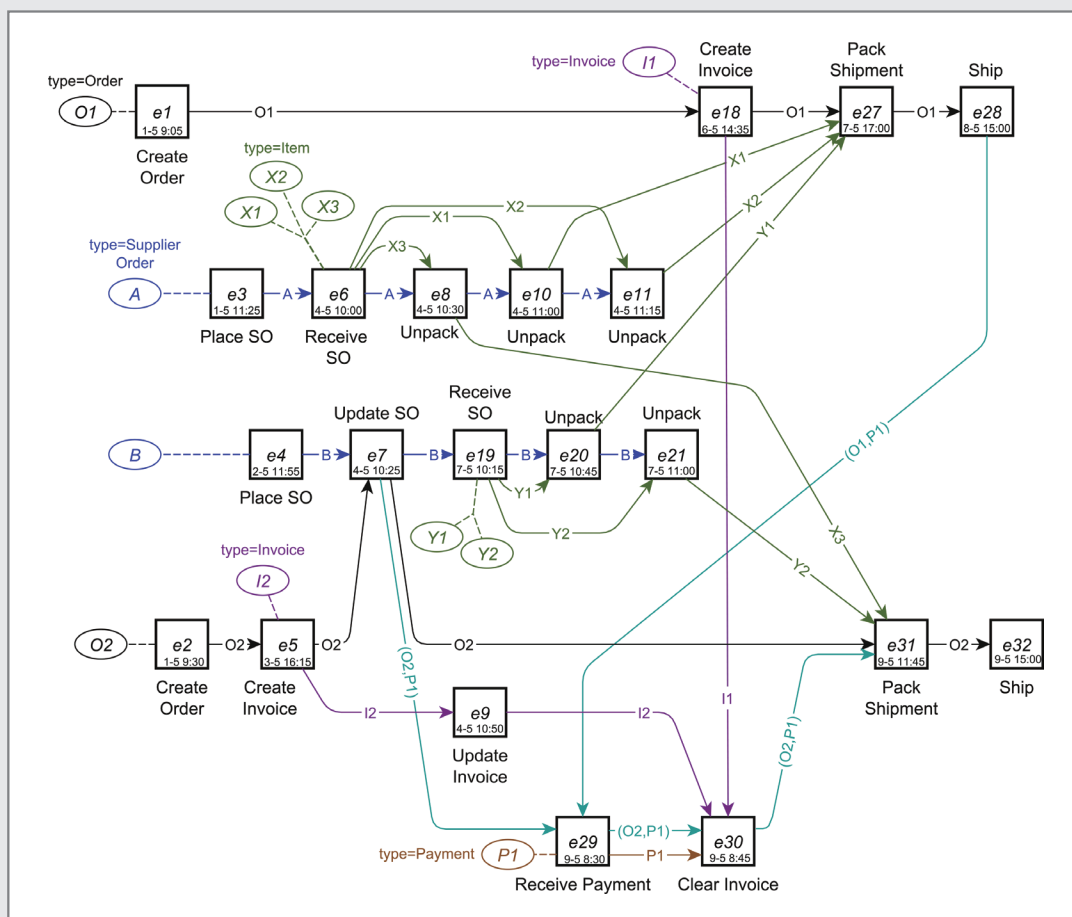
は下回ると放出になります。シーズン開始時点では放出閾値が低いことから、放出はほとんど起こらないが、契約終了が近づくにつれて放出閾値が高くなり、契約終了前に必ず放出するという結果になります。これは、契約延長がない選手は契約終了前に移籍させて移籍金を回収するという現実を反映しています。

後藤研究室では、J1 リーグのクラブと共同研究を実施しており、実際のクラブのデータから選手の契約価値を計算しています。計算結果はクラブが想定する値と大きな乖離はないというフィードバックを得ています。今後はさらなるモデルの改良に取り組み、選手の適正な金銭的価値を追求していきます。

■ 今後の展望

プロセスクラスタリングによるサッカー攻撃戦術の評価は、他に例をみない非常にユニークな研究ですが、今後の展開次第では、サッカーの戦術分析を変えてしまうような可能性も持っていると考えています。現在取り組んでいる発展として、ボール保持者以外のプロセスを統合した分析があります。この分析には、オブジェクト指向プロセスマイニング (Object-Centric Process Mining; OCPM) という手法を用います。

【図7】が OCPM のイメージを表しています。O1 と O2 の 2 つの注文と、A と B の 2 つの商品、I1 と I2 の 2 つの請求書、P1 の 1 つの支払のプロセスが関連付けられた統合プロセスになっていることが分かります。これをサッカーの攻撃プロセスに応用すれば、ボール保持者の動きとボール保持者以外の動きを関連



【図 7】 OCPM のイメージ図¹⁾。 O1 と O2 の 2 つの注文と、 A と B の 2 つの商品、 I1 と I2 の 2 つの請求書、 P1 の 1 つの支払のプロセスが関連付けられた統合プロセスになっている。

付けて分析することができるため、戦術分析の深度を飛躍的に向上させることが可能となります。

金融工学的アプローチによるサッカー選手の価値評価もまた、非常にユニークです。本稿で紹介した研究はベーシックなモデルですが、現在は多くの方向性の拡張に取り組んでいます。

まずは、移籍期間の考慮です。サッカーの移籍は限られた期間しか認められておらず、通常は夏と冬の 2 回になります。当然ながら移籍期間外は移籍が認められないため、【図 6】の放出閾値も変化することになります。次に、契約延長の考慮です。主力選手や若手選手には、契約終了 1 年前に延長交渉がなされるのが通例です。契約延長が決まっていれば、【図 5】のような時間経過による急激な価値の低下がなくなり、より現実的な契約価値の計算が可能になります。

以上のように、後藤研究室では今後も継続して、サッカーを科学的に分析する最先端の研究を進めていきます。本稿をとおして、数学とデータを使ってスポーツを科学することの面白さを感じてもらえたら、幸いです。

参考文献

- 1) Fahland, D.: "Process Mining over Multiple Behavioral Dimensions with Event Knowledge Graphs," In van der Aalst, W. M. P. and Carmona, J. (eds.): *Process Mining Handbook*, Springer, Cham, 274–319, 2022.
- 2) Kudo, Y., Shimoshimizu, M., Goto, M., Williams, M., Noguchi, N., Sasaki, S. and Takai, A.: "An Option Pricing Framework for Valuation of Football Players: Transfer Offers and Sellouts," *Proceedings of ACMMSA 2023*, 412–426, 2024.
- 3) Tunaru, R., Clark, E., Viney, H.: "An Option Pricing Framework for Valuation of Football Players," *Review of Financial Economics*, **14**, 281–295, 2005.