

A Causality-Inspired Adjusted Plus-Minus Model for Player Evaluation in Team Sports

스포츠 딥러닝/머신러닝 논문읽기 세션



오늘 발표 흐름

1

문제 정의

왜 선수 가치 추정이 어려운가

2

기존 **APM**

무엇을 잘하고 무엇을 못하는가

3

CAPM 방법론

잠재 교란을 어떻게 보정하는가

4

실험 결과

무엇을 보여주는가

5

활용 가능성

야구/축구 확장 아이디어

왜 팀 스포츠에서 선수 개인 가치 추정이 어려운가



동료 효과

선수는 항상 동료와 함께 뛰다. 팀 결과는 개인이 아닌 집합적 산물이다.



감독의 편향된 선택

감독의 선발 결정 자체가 이미 편향된 의사결정일 수 있다.



상대·맥락 혼재

상대 수준, 홈/원정, 경기 상황에 따라 같은 선수도 다른 역할을 맡는다.



Adjusted Plus-Minus(APM)이란 무엇인가

APM의 구조

각 선수에 기여도 계수 β_l 를 부여하는 **penalized linear regression** 모델

입력

Line-up 벡터 (우리 팀 +1, 상대 팀 -1)

목표

경기/세그먼트 점수차
(**score differential**)
설명

정규화

Ridge regularization $\lambda \| \beta \|_2^2$ (**collinearity** 완화)

APM 목적함수

$$\beta^* = \arg \min_{\beta} \sum_i \left[y_i - \sum_{\ell: t_{\ell}=u_i} \beta_{\ell} a_{i,\ell} - \sum_{\ell: t_{\ell}=v_i} \beta_{\ell} a_{i,\ell} \right]^2 + \lambda \| \beta \|_2^2$$

(Ridge-penalized least squares on lineup segment)

출력된 회귀계수 β_l 를 선수 가치처럼 해석

관측 상관과 인과효과는 다르다

U (숨은 교란)

감독 전술 / 상대 전력
수준

Line-up 선택

Treatment — U의
영향을 받아 결정됨

경기 결과 Y

Outcome — U와 Line-up 모두의 영향을 받음

$U \rightarrow \text{Line-up}$, $U \rightarrow Y$, $\text{Line-up} \rightarrow Y$: 전형적인
confounding 구조

APM의 인과적 한계

선택 편향

감독은 쉬운 경기와 어려운 경기에서 다른 선수를 기
용한다.

숨은 교란(Latent U)

특정 선수는 약한 상대를 상대로 더 자주 출전할 수 있
다.

혼재된 계수

β 는 선수 실력과 선택 편향이 섞인 값일 수 있다.

선수 가치를 **Counterfactual Effect**로 정의하기

이 논문의 핵심 기여: 선수 가치를 "그 선수를 0가치 대체 선수로 바꿨을 때 기대 점수차가 얼마나 변하는가"로 재정의한다.

Player Value (Counterfactual Estimand)

$$\Lambda_\ell := \mathbb{E}[Y; do(a_\ell = 1, a_\emptyset = 0, u = t_\ell)] - \mathbb{E}[Y; do(a_\ell = 0, a_\emptyset = 1, u = t_\ell)]$$

$\emptyset$ (replacement player)

0가치 대체 선수. 선수 ℓ 을 빼고 이 선수를 넣었을 때의 결과를 상정

do() 연산자

실제 관측값이 아닌 개입 기반의 기대값

APM β^* vs CAPM Λ_ℓ

논문은 β^* 가 아닌 Λ_ℓ 를 추정하는 것을 목표로 한다.

Confounding, Backdoor Adjustment, Positivity

조건부 기대값

$$\mu_\ell(w, a_\ell, a_\emptyset) = \mathbb{E}[Y \mid W = w, A_\ell = a_\ell, A_\emptyset = a_\emptyset, U = t_\ell]$$

Backdoor Identification

$$\Lambda_\ell = \int p(w) [\mu_\ell(w, 1, 0) - \mu_\ell(w, 0, 1)] dw$$

직관: "같은 맥락 w 에서 선수만 바꿔본 뒤 평균낸다."

세 가지 핵심 개념

Confounding

Line-up과 결과 모두에 영향을 주는 숨은 변수 W 의 존재.

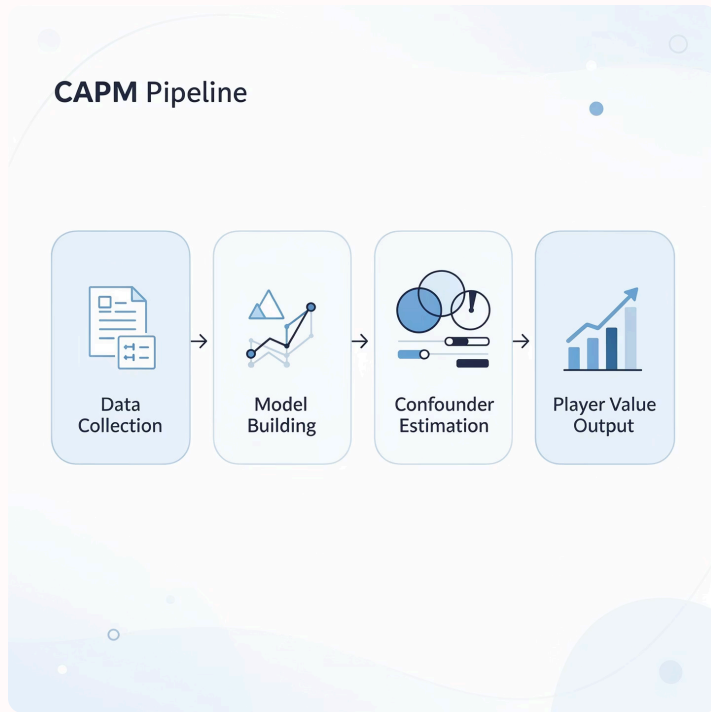
Backdoor Adjustment

같은 맥락 안에서 선수만 바꾼 차이를 평균내는 방법.
현실 문제는 W 를 직접 못 본다는 것.

Positivity

어떤 맥락에서도 비교하려는 출전 조합이 완전히 불가능하면 안 된다.

CAPM 파이프라인 한눈에 보기



Algorithm 1 요약

- ① 라인업 데이터 보고 잠재 구조 학습 (**factor model**)
- ② 경기별 **substitute confounder** 생성
- ③ **confounder**를 포함한 결과 모델로 회귀
- ④ 선수별 $\Lambda \ell$ 를 얻는다

최종 해석

$\Lambda \ell \approx \hat{\beta} \ell$ APM보다 더 **causal-like**한 player value 추정

Poisson Factorization으로 Line-up의 잠재 구조 보기

Line-up Factor Model

$$q_u \sim \text{Gamma}^K(\cdot), \quad \theta_\ell \sim \text{Gamma}^K(\cdot)$$

$$a_{i,\ell} \sim \text{Poisson}(\theta_\ell \cdot q_{v_i}) \quad \text{if } s_\ell = u_i$$

$$a_{i,\ell} \sim \text{Poisson}(\theta_\ell \cdot q_{u_i}) \quad \text{if } s_\ell = v_i$$

$$a_{i,\ell} = 0 \quad \text{otherwise}$$

$\theta_\ell \cdot q_v$: 선수 특성과 상대팀 특성의 **affinity**

모델의 직관

잠재벡터 θ_ℓ

각 선수의 비음수 잠재 특성 벡터 (K차원)

잠재벡터 q_u

각 팀의 비음수 잠재 특성 벡터 (K차원)

Bayesian NMF 계열

결과 점수를 쓰지 않고 **line-up**만으로 학습. 감독 성향·전술 대응의 흔적을 잠재요인으로 압축.

잠재 교란을 어떻게 추정하는가

Substitute Confounder

$$\hat{z}_i = (\hat{q}_{u_i}, \hat{q}_{v_i})$$

Mean-field Variational Family

$$h(\theta, q) = \prod_{\ell, k} h(\theta_{\ell k} \mid \gamma_{\ell k}^{sh}, \gamma_{\ell k}^{rt}) \times \prod_{v, k} h(q_{vk} \mid \lambda_{vk}^{sh}, \lambda_{vk}^{rt})$$

Posterior Geometric Expectations

$$\hat{\theta}_{\ell k} = \frac{\exp\{\Psi(\gamma_{\ell k}^{sh})\}}{\gamma_{\ell k}^{rt}}, \quad \hat{q}_{vk} = \frac{\exp\{\Psi(\lambda_{vk}^{sh})\}}{\lambda_{vk}^{rt}}$$

VI를 쓰는 이유와 검증

Scalable Posterior

Exact posterior 계산이 어렵기 때문에 variational inference로 근사.

K 선택

5-fold cross-validation, 평가 지표: held-out log-likelihood

Posterior Predictive Check

p-value가 0.5에 가까움으로 보고 — 모델 적합도 확인

점수차 회귀에 **Confounder**를 넣으면 무엇이 달라지나

Outcome Model

$$\mathbb{E}[Y \mid A = a_i, Z_i = z_i] = \left[\sum_{\ell} \mathbf{1}(t_{\ell} = u_i) a_{i,\ell} \beta_{\ell} + \gamma \cdot q_{v_i} \right] - \left[\sum_{\ell} \mathbf{1}(t_{\ell} = v_i) a_{i,\ell} \beta_{\ell} + \gamma \cdot q_{u_i} \right]$$

MLE Objective

$$(\hat{\beta}, \hat{\gamma}) = \arg \max \sum_i \log p(y_i \mid a_i, \hat{q}_{u_i}, \hat{q}_{v_i}; \beta, \gamma)$$

β_{ℓ} (선수 계수)

선수 고유 기여. 최종 해석: $\Delta \ell \approx \hat{\beta}_{\ell}$

$\gamma \cdot q$ (confounder 항)

잠재 전술/선택 구조의 영향을 흡수. APM에 없던 핵심 추가 항.

구현

Gaussian response + Ridge regularization. APM 파이프라인에 연결하기 쉬운 구조.

Reconstructed Causes는 왜 필요한가

Reconstructed Cause

$$\hat{a}_{i\ell} := \mathbb{E}[a_{i\ell} \mid z_i] = \sum_k z_{iuk} \theta_{\ell k}$$

Outcome Model with Reconstructed Causes

$$\mathbb{E}[Y \mid A, Z] = \sum_{\ell} \mathbf{1}[t_{\ell} = u_i] (\beta_{\ell} a_{i\ell} + \gamma_{\ell} \hat{a}_{i\ell}) - \sum_{\ell} \mathbf{1}[t_{\ell} = v_i] (\beta_{\ell} a_{i\ell} + \gamma_{\ell} \hat{a}_{i\ell})$$

아이디어의 직관

실제 출전 **$a_{i\ell}$**

관측된 실제 출전 여부

재구성 출전 **$\hat{a}_{i\ell}$**

잠재 구조가 설명하는 "예상 출전" — 구조적 부분을 분리

실험 결과

저자들은 **reconstructed causes** 버전이 **direct substitute confounder**보다 자주 더 잘 된다고 언급

합성 데이터에서는 무엇을 검증했나

Synthetic Setup (Soccer-like)

항목	값
팀 수	20
선수 수	~500
경기 수	380
라인업 크기	11명
팀당 top player	1명
시뮬레이션 횟수	100회

Outcome generation: Appendix C, Eq. 33

Confounding Regimes

수준	σ	σ_{TP}	σ_{γ}
Zero	0.1	0.1	0.01
Low	0.2	0.2	0.2
Medium	0.2	0.2	0.5
High	0.1	0.1	0.5

Figure 1 / Figure 4 메시지: Confounding이 커질수록 CAPM이 APM보다 더 잘 복원한다. 평가 지표: Spearman(rank_GT, rank_inf) 및 MSE(β_{GT} , β_{inf}). 각 점은 100회 시뮬레이션의 median, 막대는 quartiles.

실제 축구와 농구 데이터에서의 평가

Table 1 – 데이터셋 요약

리그	팀	경기	시즌
EPL	20	380	2014-16
Serie A	20	380	2014-16
La Liga	20	380	2014-16
Bundesliga	18	306	2014-16
NBA	30	~1300/ 시즌	2014-18

축구 선수 수: M = 417~515 | FIFA rating: 9,000+ reviewers, 18,000+ players 기반

평가 방법

평가 1: 랭킹 일치도

FIFA ranking과의 Spearman rank correlation

평가 2: 예측 MAE

Held-out 경기 점수차 MAE (unshifted / shifted)

데이터 차이

축구: 선발 line-up만 사용 (교체 정보 누락) 농구: substitution segment 단위로 더 세밀

논문 결과를 어떻게 읽어야 하나

Figure 2 – 랭킹 일치도

CAPM이 2/3 soccer datasets에서 FIFA ranking과 더 잘 맞음 (4 leagues × 3 seasons = 12 datasets)

Table 2 – Unshifted MAE

22개 datasets 중 18/22에서 CAPM 우위 EPL 2014: 1.36→1.31 | Serie A 2014: 1.26→1.21 La Liga 2015: 1.27→1.21 | NBA East 2017: 2.373→2.361

Table 3 – Shifted MAE

22개 datasets 중 17/22에서 CAPM 우위 EPL 2014: 1.501→1.460 | Serie A 2016: 1.459→1.392 La Liga 2016: 1.447→1.334 | NBA West 2015: 2.283→2.269

이 논문의 의의와 한계

의의

선수 가치의 정의를 회귀계수 → **counterfactual effect**로 전환

APM 위에 **latent confounding correction**를 얹는 현실적 구조 제안

실제 데이터에서도 전반적 개선 확인

한계

- **Single-player confounder**는 잡지 못함
- **Factor model**이 단순하고 **collinearity** 문제 잔존
- **Posterior mean approximation bias**
- 시간 구조, 교체, 세밀한 상호 작용 미반영



야구/축구 후속 연구 아이디어

축구 확장

교체 타이밍·포메이션의 인과효과를 분석하고, **xG·pressing**을 **outcome**으로 설정한다.

tracking-event 데이터를 결합해 선수 조합 효과와 감독 효과를 분리

야구 확장

타순·수비 시프트·볼펜 투입의 **causal effect**를 추정하고, 투구 결과를 **outcome**으로 두기.
PITCHf/x 데이터를 활용해 투구 수준의 인과 구조를 더 정교하게 반영

일반화

농구·하키 등 다른 팀 스포츠로 확장하고, **observed confounder**를 추가. 선수 조합 상호작용 모델링으로 복잡한 팀 효과까지 포착.