日本オペレーションズ・リサーチ学会
2022年 春季研究発表会

数理最適化によるテストの配点決定

	東京理科大学	*相原麟太郎	AIHARA Rintaro
02203940	東京理科大学	鮏川矩義	SUKEGAWA Noriyoshi
01012660	東京理科大学	池辺淑子	IKEBE Yoshiko
05001388	Classi 株式会社	石井康貴	ISHII Yasutaka
05001379	Classi 株式会社	廣田正之	HIROTA Masayuki

1. はじめに

Classi 株式会社は日本の高等学校向けに教育プラットフォームサービス「Classi」を提供しており，その機能の一つに「Web テスト」がある．Web テストでは問題バンクから自由に問題を選択してテストを作成・配信できる．

本発表では，Web テストの実用場面で生じる配点決定の課題について議論する．配点は学生のテストの点数の分布，したがって，学生のモチベーションや成績評価に影響を与える重要な要素であるといえる．本発表では，数理最適化によって教員の配点決定を支援する方法を提案する．さらに，提案法の有効性を Classi に蓄積された実データを用いた数値実験によって検証する．

2. 設定

テストを構成する問題は所与とする．これらの問題の添え字集合を I とおく．各問題に関する十分なデータが蓄積・分析されており，項目反応モデル [1] によって各問題の難易度や識別度が推定されている状況を考える．より具体的には，

(i) テストを受験する各学生の能力値 $\theta \in \mathbb{R}$ が推定されており，また，

(ii) 各問題 $i \in I$ の特性が（項目反応曲線と呼ばれる）横軸を能力値 θ，縦軸を能力値 θ の学生が問題 i に正答する確率 $p_{i\theta} \in [0,1]$ とするグラフとして表現されている

という状況を考える．

ここでは，能力値を離散化して能力値の集合を Θ とおき，教員（= Classi の Web テストの利用者）が各能力値 $\theta \in \Theta$ に対して能力値 θ を持つ学生のテストの点数の理想値（目標値）$t_\theta \in \mathbb{R}$ を表明している状況を考える．目標は理想値からの乖離が最小となる配点を求めることである．

3. 提案モデル

配点を決定するために数理最適化モデルを用いる．各問題 $i \in I$ の配点を表す整数変数 $w_i \in \mathbb{N}$ を用意し，これらを並べた $|I|$ 次元ベクトルを $\mathbf{w}$ とおく．各問題への正答が互いに独立であると仮定する．このとき，期待値の線形性から，能力値 θ の学生の点数の期待値は

$$E_\theta(\mathbf{w}) = \sum_{i \in I} w_i p_{i\theta}$$

と $\mathbf{w}$ の線形式で表現できる．

また，分散を考慮する．これは同じ期待値でも分散に差が生じる場合があるためである．前述の独立性の仮定より，分散の線形性から，能力値 θ の学生の点数の分散は

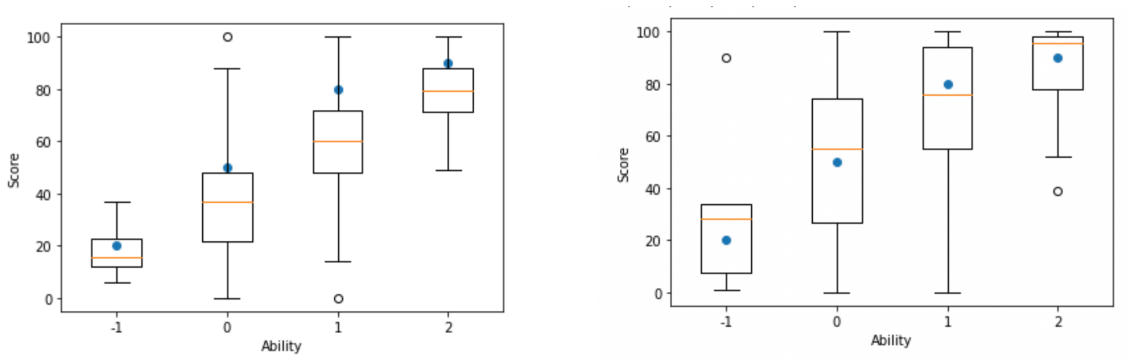
$$V_\theta(\mathbf{w}) = \sum_{i \in I} w_i^2 \, p_{i\theta}(1 - p_{i\theta})$$

と $\mathbf{w}$ の 2 次式で表現できる．ここでは，能力値に関する分散の和に上限を課すことを考える．この制約は 2 次推制約となる．

以上をまとめると，前節の目標で述べた所望の配点を求める問題は以下の数理最適化問題として定式化できる．

$$\begin{aligned}
\text{minimize} \quad & \sum_{\theta \in \Theta} (t_\theta - E_\theta(\mathbf{w}))^2 \\
\text{subject to} \quad & \sum_{\theta \in \Theta} V_\theta(\mathbf{w}) \le C, \\
& \sum_{i \in I} w_i = 100, \quad w_i \in \mathbb{N} \ (\forall i \in I).
\end{aligned}$$

ここで C は分散和の上限で，教員が定めるパラメータとする．また，今回は Classi のデータを用いて実験をするため，Classi のテストの仕様に合わせ，満点を 100 点としている．



(a) $C = 250$ の場合 (b) $C = 1000$ の場合

図 1: 実データに対する数値実験結果：横軸は能力値 $\theta \in \Theta = \{-1, 0, 1, 2\}$（$\theta$ を離散化してプロットしている），縦軸はテストの点数．各青点は能力値 θ を持つ学生に対する点数の理想値 $t_\theta \in \mathbb{R}$ を表す．

問題数が少ない場合は，商用ソルバーによってこの問題を比較的高速に解くことができる．しかし，問題数が多い場合は，変数（配点）の整数制約を連続緩和したり，目的関数の 2 乗誤差を絶対誤差に置き換えたりするなどの工夫がないと，現実的な時間で解くことができない．

4. 予備実験

この予備実験では過去に行なわれた Classi の実際の Web テストのデータを用いる．以降，同テストをテスト X，テスト X を利用した教員を教員 X と呼ぶ．テスト X は 15 問の問題から成り，受験の対象となる学生は 236 人である．

実験の流れは以下のとおりである．まず，項目反応理論に基づいて，各問題の項目反応曲線を推定した．これには Python の pyirt ライブラリを用いた．つぎに，提案モデルによってテスト X の各問題の配点を決定した．計算時間は 5 分程度であった（CPU:corei9-10900k OS:windows）．その後，学生のテスト X における各問題の（実際の）正誤の情報を用い，各能力値毎の点数の分布を確認した．なお，教員 X のテスト X における点数の理想値 $(t_\theta)_{\theta\in\Theta}$ はこちらから仮想的に与えた．

パラメータ C の値として 250 と 1000 の 2 パターンを検証した．結果を図 1 に示す．$C = 1000$ の場合に比べ，$C = 250$ の場合は各能力値において箱ひげ図が小さくなっており，生徒の点数が比較的分散していないことがわかる．これは提案モデルにおける分散和の上限制約が適切に作用したためである．一方，点数の理想値からの期待値の乖離の観点からは，$C = 1000$ の場合のほうが好ましい結果となっている．これは提案モデルにおける分散和の上限制約が緩められたためである．

5. 今後の課題

ここでは各問題への正答が互いに独立であると仮定したが，この仮定が適切でない可能性がある．適切でない場合，特に良い構造がなければ，各問題への正誤の割当の $2^{|I|}$ パターンのシナリオを考える必要があり，計算量的に困難である．また，教員が納得する配点であるかも重要である．前述の実験では，$C = 250$ の場合に比べ，$C = 1000$ の場合は非常に不均衡な（実際に用いることに抵抗感がある）配点となっていた．教員の抵抗感を排除するために，配点の範囲や大小関係に関する適切な制約を課すことも検討している．

また，ここではテスト開始前に配点を決定し，問題に付記する状況を想定したが，テスト終了後に配点を決定・修正する方法も検討されており [2]，そのような場面への数理最適化モデルの応用も考えられる．

参考文献

[1] R.K. Hambleton and H. Swaminathan: Item Response Theory: Principles and Applications (Springer, 2013).

[2] 前川眞一: 大学入試センター試験における選択科目間の得点調整について. 計測と制御, **40**(8), pp. 568–571.