مجله مدل‌سازی پیشرفته ریاضی
سال ۱۴۰۰، دوره ۱۱، شماره ۱، ص ۶۱–۶۸

شاپا چاپی: ۲۲۵۱–۸۰۸۸
DOI: 10.22055/jamm.2021.29978.1731

معیاری برای انتخاب طرح عاملی کسری دوسطحی

نبز اسمعیل‌زاده ×، شیدا رمضانی

گروه آمار، دانشگاه کردستان، سنندج، ایران

تاریخ دریافت: ۱۳۹۸/۳/۲۷ تاریخ پذیرش: ۱۳۹۹/۸/۳
دبیر مسئول: فریبرز آذرپناه

چکیده: استفاده از طرح‌های عاملی کسری در آزمایش‌هایی که شامل تعداد زیادی عامل‌اند، رایج است. انتخاب کسر مناسب مساله مهمی در پژوهش‌های طرح‌های کسری است. معیارهای مختلفی براساس دیدگاه‌های متفاوتی وجود دارند. در این مقاله معیاری براساس مینیمم‌سازی تابعی موزون از ماتریس میانگین مربعات خطای برآوردگرهای حداقل‌مربعات مدلی از پیش مشخص معرفی می‌شود. در طرح‌های کسری دوسطحی، این معیار برای تابع وزن یکنواخت محاسبه و نشان داده می‌شود که با معیار $D-$ بهینه شناخته شده معادل است. در پایان با ذکر مثال‌هایی روش تشریح می‌شود.

واژه‌های کلیدی: طرح عاملی کسری، طرح بهینه، طرح دوسطحی.

رده‌بندی ریاضی: 62k05; 62k15

۱ مقدمه

بسیاری از آزمایش‌ها شامل مطالعه‌ی اثر بیش از یک عامل روی متغیر پاسخ هستند. طرح‌های عاملی برای شناسایی عامل‌های مؤثر در این نوع آزمایش‌ها کارا هستند. در یک طرح عاملی کامل در هر تکرار آزمایش تمام ترکیب‌های تیماری بررسی می‌شوند. یک طرح عاملی کامل امکان آزمودن همه‌ی اثرات اصلی و متقابل میان عامل‌ها را فراهم می‌کند. متاسفانه، در این طرح‌ها تعداد کل آزمایش‌ها با افزایش عامل‌ها سریع افزایش می‌یابد و امکان اجرای همه‌ی آزمایش‌ها را در عمل با مشکل مواجه می‌کند. در این حالت استفاده از طرح‌های عاملی کسری رایج است.

فرض کنیم N تعداد اجراهای یک طرح عاملی و $n < N$ تعداد اجراهای یک طرح کسری باشد. تعداد $\binom{N}{n}$ انتخاب برای تعیین یک طرح عاملی کسری وجود دارد. سؤال این است: کدام ترکیبات تیماری باید انتخاب شوند؟ به‌طور کلی در انتخاب یک طرح عاملی کسری دو راهبرد وجود دارد. در اولی اثرات متقابل مرتبه بالا را ناچیز فرض می‌کنند و کسری انتخاب می‌شود که توانایی برآورد اثرات مرتبه پایین را داشته باشد. این راهبرد متکی بر اصل سلسله مراتبی اثرات است. طبق این اصل اثرات مرتبه پایین اهمیت بیشتری نسبت به اثرات مراتب بالاتر دارند و اثرات هم مرتبه از اهمیت یکسانی برخوردارند[۱۱] . رایج‌ترین معیار بر اساس این اصل، معیار ماکزیمم وضوح است که توسط باکس و هانتر [۱، ۲] پیشنهاد شد. همچنین می‌توان به معیارهای وضوح تعمیم یافته و کمترین انحراف اشاره کرد.

×نویسنده مسئول مقاله
رایانامه: n.esmailzadeh@uok.ac.ir (N. Esmailzadeh), shaydaramezani67@gmail.com (S. Ramezani)

در راهبرد دوم، زیرمجموعه‌ای از اثرات عاملی را که به نظر معنی‌دارند، مشخص می‌کنند. این زیرمجموعه را با $R_۰$ نشان می‌دهیم. سپس براساس این مجموعه، یک مدل رگرسیون کلی مانند $E[y] = \phi(\mathrm{u}, \boldsymbol{\theta})$ در نظر می‌گیرند، که در آن $\phi(\mathrm{u}, \boldsymbol{\theta})$ تابعی از سطوح عامل‌ها و اثرات عاملی است. براساس این مدل از پیش مشخص، معیارهای بهینگی مختلفی برای انتخاب کسر وجود دارند. هدف پیدا کردن طرحی است که برآوردهای مطلوبی برای پارامترهای مجهول $\boldsymbol{\theta}$ ارایه کند.

متداول‌ترین معیارهای بهینگی مبتنی بر مدل معیارهای $D-$ بهینگی [†]، $A-$ بهینگی و $E-$ بهینگی هستند که بر اساس اندازه‌ای از ماتریس کوواریانس برآوردگرهای حداقل‌مربعات تعریف می‌شوند [۷]. از این معیارها برای ساخت طرح‌های بهینه در مدل‌هایی شامل اثرات اصلی [۳] یا اثرات اصلی و تعدادی از اثرات متقابل دو عاملی [۹] استفاده شده است.

اگر مدل از پیش مشخص درست نباشد، برآوردگرهای حداقل‌مربعات ناریب نیستند. کاهش اریبی برآوردگرها موضوع مهمی است که در معیارهای بهینگی ذکرشده در نظر گرفته نمی‌شود. برای لحاظ کردن این موضوع، معیارهای بهینگی مینیماکس بر اساس اندازه‌ای از ماتریس میانگین مربعات خطای برآوردگرهای حداقل‌مربعات تعریف شده‌اند. معیار $D-$ بهینگی مینیماکس و $A-$ بهینگی مینیماکس به‌ترتیب ماکزیمم دترمینان و ماکزیمم اثر ماتریس میانگین مربعات خطا را مینیمم می‌سازند. معیار $D-$ بهینگی مینیماکس توسط ویلموت و ژو [۱۰] برای ساخت طرح‌های عاملی کسری دوسطحی بهینه پیشنهاد شد. لین و ژو [۶] آن را برای طرح‌های عاملی کسری سه‌سطحی و سطوح آمیخته بسط دادند. معیار $A-$ بهینگی مینیماکس نیز در [۱۲] پیشنهاد شده است.

اگر مجموعه $R_۰$ درست تعیین نشود، میانگین مربعات خطای برآوردگرهای حداقل‌مربعات به پارامترهایی وابسته است که در آن مجموعه قرار ندارند. معیارهای بهینگی مینماکس این وابستگی را با ماکزیمم‌سازی روی این پارامترها رفع می‌کنند. سپس طرحی که کمیت حاصل را مینیمم می‌کند به‌عنوان طرح بهینه انتخاب می‌شود. مانند همه قواعد مینیماکس، این معیارها نیز در مواردی ممکن است کاملا غیرمنطقی باشند [۵]. همچنین ممکن است بعضی اوقات اطلاعاتی در مورد اهمیت و شانس معنی‌داری اثرات عاملی موجود باشد. معیارهای مینیماکس توانایی داخل کردن چنین اطلاعاتی را ندارند. در این مقاله این اطلاعات را در قالب تابعی وزنی از اثرات وارد مساله می‌کنیم. با متوسط‌گیری وزنی از تابعی از میانگین مربعات خطا روی پارامترهایی که در $R_۰$ نیستند و سپس مینیمم کردن کمیت حاصل، دنبال طرح بهینه می‌گردیم. ادامه مقاله بدین صورت سازماندهی شده است. در بخش ۲ نمادگذاری‌ها و مدل آماری مورد نیاز ارایه می‌شوند. در بخش ۳ معیار WI–بهینه معرفی می‌شود. همچنین محاسبه آن برای طرح‌های دوسطحی همراه با مثال ارایه می‌شود.

۲ نمادگذاری و مدل آماری

فرض کنیم $F_۱$، $F_۲$، ...، F_k، k عامل دوسطحی در یک طرح عاملی باشند. سطوح این عوامل را با $+۱$ و -۱ نشان می‌دهیم. در این طرح تعداد اجراها برابر $N = ۲^k$ است. ماتریس طرح را با $\mathrm{D} = [\mathrm{d}_۱, ..., \mathrm{d}_N]^T$ نشان می‌دهیم که برای $i = ۱, ..., N$، سطر $\mathrm{d}_i^T = (d_{i۱}, ..., d_{ik})$ i–ام ماتریس D است. مجموعه‌ی $\mathcal{D} = \{\mathrm{d}_۱, \ldots, \mathrm{d}_N\}$ را فضای طرح می‌نامیم. اعضای $\mathcal{D}$ را نقاط طرح می‌گوییم. در این طرح، ستون متناظر اثرات متقابل از ضرب درایه‌بدرایه ستون‌های تشکیل دهنده آن اثر حاصل می‌شوند.

برای یک طرح عاملی کامل ۲^k، مدل خطی

$$\mathrm{y} = \mathrm{U}\boldsymbol{\theta} + \boldsymbol{\epsilon}, \tag{۱.۲}$$

را در نظر می‌گیریم که در آن، y بردار متغیر پاسخ، $\mathrm{U} = (۱, \mathrm{u}_۱, \ldots, \mathrm{u}_p)$ ماتریس مدل و ۱ بردار ستونی از یک‌ها است. همچنین $\boldsymbol{\theta} = (\theta_۰, \theta_۱, \ldots, \theta_p)^T$ بردار پارامترهای مجهول متناظر با اثرات عاملی موجود در مدل است. $\boldsymbol{\epsilon}$ بردار خطاهای تصادفی با بردار میانگین $\mathbf{0}$ و ماتریس کوواریانس $\sigma^۲\mathrm{I}$ است که I ماتریس همانی از مرتبه مناسب می‌باشد. واضح است که اگر θ_i متناظر با یک اثر اصلی باشد با ستون متناظر آن در ماتریس D برابر است. اگر همه اثرات عاملی در مدل حضور داشته باشند، آنگاه $p = ۲^k - ۱$ است.

در این مقاله $R_۰$ را مجموعه‌ای شامل همه‌ی اثرات اصلی و تعدادی اثر متقابل در نظر می‌گیریم. فرض می‌کنیم تعداد اعضای این مجموعه برابر m است. بر اساس اثراتی که در مجموعه‌ی $R_۰$ حضور دارند، ماتریس U را به‌صورت $\mathrm{U} = (\mathrm{U}_۱, \mathrm{U}_۲)$ افراز می‌کنیم که در آن $\mathrm{U}_۱ = (۱, \mathrm{u}_۱, \ldots, \mathrm{u}_m)$ و $\mathrm{U}_۲ = (\mathrm{u}_{m+۱}, \ldots, \mathrm{u}_p)$. همچنین به‌طور متناظر بردار $\boldsymbol{\theta}$ نیز به دو بردار $\boldsymbol{\theta}_۱ = (\theta_۰, \theta_۱, \ldots, \theta_m)^T$ و $\boldsymbol{\theta}_۲ = (\theta_{m+۱}, \ldots, \theta_p)^T$ افراز می‌شود.

تحت این شرایط، مدل کامل (۱.۲) به مدل $\mathrm{y} = \mathrm{U}_۱\boldsymbol{\theta}_۱ + \boldsymbol{\epsilon}$ کاهش می‌یابد. یک طرح عاملی کسری با n اجرا، انتخاب n عضو از مجموعه $\mathcal{D}$ است. مدل متناظر با طرح عاملی کسری را به‌صورت

$$\mathrm{y} = \mathrm{Z}_۱\boldsymbol{\theta}_۱ + \boldsymbol{\epsilon} \tag{۲.۲}$$

می‌نویسیم که در آن $\mathrm{Z}_۱ = (۱, \mathrm{z}_۱, \ldots, \mathrm{z}_m)$ ماتریس مدل $n \times (m+۱)$، برای طرح عاملی کسری است. $\mathrm{z}_۱, \ldots, \mathrm{z}_m$ ستون‌های ماتریس $\mathrm{Z}_۱$ متناظر با اثرات موجود در $R_۰$ هستند. به‌منظور روشن شدن جزییات، مثال زیر را بیان می‌کنیم.

[†] D-optimality

مثال ۱.۲. یک طرح 2^3 با عامل‌های F_1،F_2 و F_3 را در نظر می‌گیریم. ماتریس مدل کامل به‌صورت زیر است:

$$\mathrm{U} = \begin{pmatrix} 1 & -1 & -1 & -1 & +1 & +1 & +1 & -1 \\ 1 & +1 & -1 & -1 & -1 & -1 & +1 & +1 \\ 1 & -1 & +1 & -1 & -1 & +1 & -1 & +1 \\ 1 & +1 & +1 & -1 & +1 & -1 & -1 & -1 \\ 1 & -1 & -1 & +1 & +1 & -1 & -1 & +1 \\ 1 & +1 & -1 & +1 & -1 & +1 & -1 & -1 \\ 1 & -1 & +1 & +1 & -1 & -1 & +1 & -1 \\ 1 & +1 & +1 & +1 & +1 & +1 & +1 & +1 \end{pmatrix}$$

در این مثال بردار $\boldsymbol{\theta} = (\theta_0, \theta_1, \dots, \theta_7)^T$. فرض می‌کنیم R_0 شامل همه‌ی اثرات اصلی و اثرات متقابل دو عاملی بین عامل‌های F_1 و F_2 و همچنین F_1 و F_3 است؛ یعنی

$$R_0 = \{F_1, F_2, F_3, F_1F_2, F_1F_3\}, \quad m = 5.$$

ماتریس U_1 از ستون‌های ۱ تا ۶ ماتریس U تشکیل می‌شود. اگر $n = 4$ و چهار اجرای اول انتخاب شوند، آنگاه Z_1 ماتریس 6×4 متشکل از سطرهای ۱ تا ۴ ماتریس U_1 است. همچنین $\boldsymbol{\theta}_1 = (\theta_0, \theta_1, \dots, \theta_5)^T$.

اگر اثرات معنی‌دار دیگری وجود داشته باشند که به اشتباه در R_0 وارد نشوند، یعنی R_0 به درستی تعیین نشود، آنگاه مدل درست را می‌توان به‌صورت

$$E(y_i) = \theta_0 + \sum_{j=1}^{m} \theta_j u_{ij} + f(u_{i1}, \dots, u_{im}, \theta),\ i = 1, \dots, n \tag{۳.۲}$$

نوشت [۶] که در آن f تابع دوری [‡] نامیده می‌شود. به پیروی از ویلمت و ژو [۱۰]، کلاس $\mathcal{F}$ متشکل از همه‌ی توابع دوری را در نظر می‌گیریم که در آن f در دو شرط زیر صدق می‌کند:

شرط یک) تابع f و $\mathrm{u}_0, \mathrm{u}_1, \dots, \mathrm{u}_m$ روی فضای طرح $\mathcal{D}$ متعامداند؛ یعنی،

$$\sum_{i=1}^{N} f(u_{i1}, \dots, u_{im}) u_{ij} = \sum_{i=1}^{N} f(\mathrm{d}_i) u_{ij} = 0,\ j = 0, 1, \dots, m,$$

که در آن برای $i = 1, \dots, N$، $u_{i0} = 1$ و u_{ij}، $i-$امین عنصر ستون u_j است.

شرط دو) تابع f روی فضای طرح $\mathcal{D}$ کراندار است؛ یعنی، برای یک $\alpha \geq 0$

$$\frac{1}{N} \sum_{i=1}^{N} f^2(u_{i1}, \dots, u_{im}) = \frac{1}{N} \sum_{i=1}^{N} f^2(\mathrm{d}_i) \leq \alpha^2.$$

لم زیر همه‌ی توابع موجود در کلاس $\mathcal{F}$ را نشان می‌دهد.

لم ۲.۲. [۱۰] هر تابع $f \in \mathcal{F}$ که در شرط یک و دو صدق کند، به‌صورت زیر است:

$$f(\mathrm{d}_i) = \mathrm{z}_2^T(\mathrm{d}_i)\boldsymbol{\theta}_2, \quad i = 1, \dots, N,$$

که در آن $i = 1, \dots, N$، $\mathrm{z}_2(\mathrm{d}_i) = (u_{i,m+1}, \dots, u_{i,p})^T$ و $\|\boldsymbol{\theta}_2\| \leq \alpha$ است.

لم ۲.۲ نشان می‌دهد تابع f در (۳.۲) ترکیب خطی از تمام اثراتی است که در R_0 نیستند. بزرگی f به‌وسیله‌ی کران داده شده در شرط دو کنترل می‌شود. اگر مدل (۲.۲) درست نباشد برآوردگر حداقل‌مربعات $\boldsymbol{\theta}_1$ اریب است که تابع f مقدار این اریبی را تعیین می‌کند. بنابراین از لم ۲.۲، صورت ماتریسی مدل کامل متناظر با طرح عاملی کسری عبارت است از:

$$\mathrm{y} = \mathrm{Z}_1\boldsymbol{\theta}_1 + \mathrm{Z}_2\boldsymbol{\theta}_2 + \boldsymbol{\epsilon},$$

که $\boldsymbol{\theta}_2$ بردار اثراتی است که در R_0 نیستند و Z_2 ماتریس مدل متناظر $\boldsymbol{\theta}_2$ است. سطرهای Z_2 از بین سطرهای U_2 انتخاب می‌شوند.

برآوردگر حداقل‌مربعات $\boldsymbol{\theta}_1$ در مدل (۲.۲) برابر $\hat{\boldsymbol{\theta}}_1 = (\mathrm{Z}_1^T\mathrm{Z}_1)^{-1}\mathrm{Z}_1^T\mathrm{y}$ و واریانس و اریبی $\hat{\boldsymbol{\theta}}_1$ به‌ترتیب برابر است با

[‡] Departure function

$$cov(\hat{\boldsymbol{\theta}}_۱) = \sigma^۲(\mathrm{Z}_۱^T\mathrm{Z}_۱)^{-۱},$$

$$Bias(\hat{\boldsymbol{\theta}}_۱) = E(\hat{\boldsymbol{\theta}}_۱) - \boldsymbol{\theta}_۱ = (\mathrm{Z}_۱^T\mathrm{Z}_۱)^{-۱}\mathrm{Z}_۱^T\mathrm{Z}_۲\boldsymbol{\theta}_۲.$$

بنابراین ماتریس میانگین مربعات خطا عبارت است از:

$$\begin{aligned} MSE(\hat{\boldsymbol{\theta}}_۱, \mathrm{Z}_۱, \boldsymbol{\theta}_۲) &= cov(\hat{\boldsymbol{\theta}}_۱) + Bias(\hat{\boldsymbol{\theta}}_۱)Bias(\hat{\boldsymbol{\theta}}_۱)^T \\ &= \sigma^۲(\mathrm{Z}_۱^T\mathrm{Z}_۱)^{-۱} + (\mathrm{Z}_۱^T\mathrm{Z}_۱)^{-۱}\mathrm{Z}_۱^T\mathrm{Z}_۲\boldsymbol{\theta}_۲{\boldsymbol{\theta}_۲}^T\mathrm{Z}_۲^T\mathrm{Z}_۱(\mathrm{Z}_۱^T\mathrm{Z}_۱)^{-۱}. \end{aligned}$$

می‌توان نشان داد [۸] :

$$det\big(MSE(\hat{\boldsymbol{\theta}}_۱, \mathrm{Z}_۱, \boldsymbol{\theta}_۲)\big) = \sigma^{۲(m+۱)}det\big((\mathrm{Z}_۱^T\mathrm{Z}_۱)^{-۱}\big)\big(۱ + \frac{۱}{\sigma^۲}{\boldsymbol{\theta}_۲}^T\mathrm{A}\boldsymbol{\theta}_۲\big) \tag{۴.۲}$$

که $\mathrm{A} = \mathrm{Z}_۲^T\mathrm{Z}_۱(\mathrm{Z}_۱^T\mathrm{Z}_۱)^{-۱}\mathrm{Z}_۱^T\mathrm{Z}_۲$. برای یک طرح کسری دوسطحی $\mathrm{A} = N\mathrm{I} - \mathrm{Z}_۱^T\mathrm{Z}_۱$ [۱۰]. بنابراین به‌راحتی می‌توان نشان داد که $trace(\mathrm{A}) = (N-n)(m+۱)$. منظور از $trace$ اثر ماتریس است.

۳ معیار WI -بهینه

یکی از معیارهای انتخاب طرح بهینه براساس کمینه کردن کمیت $det\big(MSE(\hat{\boldsymbol{\theta}}_۱, \mathrm{Z}_۱, \boldsymbol{\theta}_۲)\big)$ در رابطه (۴.۲) است [۴، ۷]. این کمیت به مقادیر مجهول $\boldsymbol{\theta}_۲$ بستگی دارد. بنابراین، یافتن ماتریس $\mathrm{Z}_۱$ به‌طوری‌که این کمیت را به‌ازای همه مقادیر $\boldsymbol{\theta}_۲$ مینیمم کند، امکان‌پذیر نیست. متوسط وزنی آن را روی $\boldsymbol{\theta}_۲ \in \Theta_۲$ که $\Theta_۲ = \{{\boldsymbol{\theta}_۲}^T\boldsymbol{\theta}_۲ \leq \alpha^۲\}$ در نظر می‌گیریم و برای انتخاب طرح بهینه کمیت حاصل را مینیمم می‌کنیم. تابع $L(\mathrm{Z}_۱)$ را به‌صورت

$$L(\mathrm{Z}_۱) = \int\cdots\int_{\Theta_۲} g(\theta_{m+۱}, \ldots, \theta_p)det(MSE(\hat{\boldsymbol{\theta}}_۱, \mathrm{Z}_۱, \boldsymbol{\theta}_۲))d\theta_{m+۱}\ldots d\theta_p \tag{۱.۳}$$

تعریف می‌کنیم که $g(\theta_{m+۱}, \ldots, \theta_p)$ تابع وزن است. یک طرح را $WI-$ بهینه گوییم هرگاه برای یک n ثابت، تابع $L(\mathrm{Z}_۱)$ را روی همه‌ی طرح‌های $\mathrm{Z}_۱$ ممکن مینیمم سازد.

تعریف ۱.۳. طرح $\mathrm{Z}_۱^*$ را یک طرح $WI-$ بهینه می‌نامیم هرگاه

$$L(\mathrm{Z}_۱^*) = \min_{\mathrm{Z}_۱} L(\mathrm{Z}_۱).$$

مجموعه $R_۰$ شامل اثرات بالقوه مهم است و بردار $\boldsymbol{\theta}_۲$ شامل اثراتی است که غیرفعال فرض شده و متعلق به $R_۰$ نیستند. به‌عنوان حالتی خاص، فرض می‌کنیم که $g(\theta_{m+۱}, \ldots, \theta_p)$ تابعی ثابت باشد. این متناظر با حالتی است که معتقد باشیم هرکدام از اثرات $\theta_{m+۱}, \ldots, \theta_p$ دارای اهمیت یکسان و از شانس برابری برای کاندید شدن به‌عنوان اثر غیرفعال برخوردارند. همچنین در نبود اطلاعات درباره ارجحیت این اثرات این فرض معقول و عملی به نظر می‌رسد.

این مقدار ثابت اثری روی مینیمم‌سازی تابع $L(\mathrm{Z}_۱)$ در (۱.۳) ندارد. ازاین‌رو جهت سادگی محاسبات آن‌را برابر ۱ قرار می‌دهیم. پس

$$\begin{aligned} L(\mathrm{Z}_۱) &= \int_{\boldsymbol{\theta}_۲ \in \Theta_۲}\ldots\int det\big(MSE(\hat{\boldsymbol{\theta}}_۱, \mathrm{Z}_۱, \boldsymbol{\theta}_۲)\big)d\theta_{m+۱}\ldots d\theta_p \\ &= \sigma^{۲(m+۱)}det\big((\mathrm{Z}_۱^T\mathrm{Z}_۱)^{-۱}\big)\int\ldots\int\Big\{۱ + \frac{۱}{\sigma^۲}{\boldsymbol{\theta}_۲}^T\mathrm{A}\boldsymbol{\theta}_۲\big)\Big\}d\theta_{m+۱}\ldots d\theta_p \\ &= \sigma^{۲(m+۱)}det\big((\mathrm{Z}_۱^T\mathrm{Z}_۱)^{-۱}\big)\big\{C_۱(\alpha) + C_۲(\alpha, \sigma^۲)trace(A)\big\} \\ &= \sigma^{۲(m+۱)}det\big((\mathrm{Z}_۱^T\mathrm{Z}_۱)^{-۱}\big)\big\{C_۱(\alpha) + C_۲(\alpha, \sigma^۲)(N-m)(m+۱)\big\}, \end{aligned} \tag{۲.۳}$$

جدول ۱: طرح عاملی 2^4

اجرا	F_1	F_2	F_3	F_4	اجرا	F_1	F_2	F_3	F_4
۱	-1	-1	-1	-1	۹	-1	-1	-1	$+1$
۲	$+1$	-1	-1	-1	۱۰	$+1$	-1	-1	$+1$
۳	-1	$+1$	-1	-1	۱۱	-1	$+1$	-1	$+1$
۴	$+1$	$+1$	-1	-1	۱۲	$+1$	$+1$	-1	$+1$
۵	-1	-1	$+1$	-1	۱۳	-1	-1	$+1$	$+1$
۶	$+1$	-1	$+1$	-1	۱۴	$+1$	-1	$+1$	$+1$
۷	-1	$+1$	$+1$	-1	۱۵	-1	$+1$	$+1$	$+1$
۸	$+1$	$+1$	$+1$	-1	۱۶	$+1$	$+1$	$+1$	$+1$

جدول ۲: طرح‌های WI- بهینه در مثال ۲.۳

n	تعداد	شماره اجرا از جدول ۱
۸	۵۱۲	۱, ۲, ۳, ۴, ۵, ۶, ۷, ۹
۹	۱۲۸	۱, ۲, ۳, ۴, ۵, ۶, ۷, ۸, ۱۰
۱۰	۳۲	۱, ۲, ۳, ۴, ۵, ۶, ۷, ۸, ۱۰, ۱۱
۱۱	۶۴	۱, ۲, ۳, ۴, ۵, ۶, ۷, ۸, ۹, ۱۰, ۱۲
۱۲	۸	۱, ۲, ۳, ۴, ۵, ۶, ۷, ۹, ۱۰, ۱۱, ۱۲, ۱۳
۱۳	۸	۱, ۲, ۳, ۴, ۶, ۷, ۹, ۱۰, ۱۱, ۱۲, ۱۳, ۱۴
۱۴	۲۱	۱, ۲, ۳, ۴, ۵, ۶, ۷, ۸, ۹, ۱۰, ۱۱, ۱۲, ۱۳, ۱۴
۱۵	۲	۱, ۲, ۳, ۴, ۵, ۶, ۷, ۸, ۹, ۱۰, ۱۱, ۱۲, ۱۳, ۱۴, ۱۵

که در آن $C_1(\alpha)$ و $C_2(\alpha, \sigma^2)$ توابعی از α و σ هستند که به طرح بستگی ندارند. در محاسبه انتگرال (۲.۳) از رابطه

$$\boldsymbol{\theta}_2{}^T \mathrm{A} \boldsymbol{\theta}_2 = \sum_{i=m+1}^{p} \sum_{j=m+1}^{p} a_{ij}\theta_i\theta_j$$

استفاده شده است که a_{ij} درایه‌های ماتریس A هستند.

با توجه به رابطه (۲.۳)، مینیمم کردن $L(\mathrm{Z}_1)$ در رابطه (۱.۳) معادل مینیمم کردن $det\left((\mathrm{Z}_1^T\mathrm{Z}_1)^{-1}\right)$ است. پس Z_1^* طرحی است که $det\left((\mathrm{Z}_1^T\mathrm{Z}_1)^{-1}\right)$ را مینیمم می‌کند. یعنی با انتخاب تابع وزن ثابت، این طرح با طرح D -بهینه معادل است.

مثال ۲.۳. طرح 2^4 داده‌شده در جدول ۱ را در نظر می‌گیریم. فرض کنیم

$$R_0 = \{F_1, F_2, F_3, F_4, F_1F_2, F_1F_3\}, \quad N = 16, \quad m = 6.$$

برای یافتن طرح بهینه به‌وسیله‌ی الگوریتم جستجوی کامل همه‌ی $\binom{N}{n}$ انتخاب ممکن را بررسی می‌کنیم. سپس طرحی که رابطه (۲.۳) را مینیمم می‌کند، به‌عنوان طرح بهینه انتخاب می‌کنیم. برای $n = 8, \ldots 15$ به‌وسیله‌ی جستجوی کامپیوتری و برنامه‌نویسی در محیط نرم‌افزار MATLAB همه‌ی طرح‌های WI - بهینه را جستجو می‌کنیم. تعداد این طرح‌ها و نماینده‌ای از آنها براساس شماره اجراها از جدول ۱ برای هر n در جدول ۲ آورده شده‌اند. دراین جدول ستون n، تعداد ترکیبات تیماری (اجراها) مورد نظر در طرح کسری است. ستون "تعداد"، تعداد کل طرح‌های بهینه موجود برای n مفروض است. مثلا اگر $n = 8$ باشد، $\binom{16}{8} = 12870$ حالت برای انتخاب طرح کسری با ۸ اجرا وجود دارد که در بین آنها ۵۱۲ کسر براساس معیار معرفی‌شده بهینه هستند. ستون "شماره اجرا" نماینده‌ای از بین تعداد کل طرح‌های بهینه در ستون "تعداد" است که در آن شماره‌ها، شماره اجراها از ستون "اجرا" در جدول ۱ است. مثلا برای $n = 8$، یکی از طرح‌های کسری با ۸ اجرا دارای ترکیبات تیماری شماره ۷-۱ و ۹ از جدول ۱ است.

جدول ۳: طرح‌های WI– بهینه در مثال ۳.۳

n	شماره اجرا از جدول ۱
۸	۲, ۴, ۷, ۸, ۹, ۱۰, ۱۲, ۱۶
۹	۱, ۲, ۳, ۶, ۷, ۸, ۱۲, ۱۵, ۱۶
۱۰	۲, ۴, ۵, ۶, ۸, ۱۰, ۱۲, ۱۴, ۱۵, ۱۶
۱۱	۱, ۲, ۳, ۷, ۹, ۱۰, ۱۱, ۱۲, ۱۳, ۱۴, ۱۶
۱۲	۱, ۲, ۳, ۴, ۶, ۸, ۹, ۱۰, ۱۲, ۱۳, ۱۵, ۱۶
۱۳	۱, ۲, ۳, ۴, ۵, ۷, ۹, ۱۰, ۱۱, ۱۲, ۱۳, ۱۵, ۱۶
۱۴	۱, ۲, ۳, ۴, ۵, ۶, ۷, ۸, ۹, ۱۰, ۱۱, ۱۲, ۱۴, ۱۵
۱۵	۱, ۲, ۳, ۴, ۵, ۶, ۷, ۸, ۹, ۱۰, ۱۱, ۱۲, ۱۳, ۱۴, ۱۵

ذکر چند نکته درباره انتخاب تابع وزن ضروری است. اگر به طریقی مانند مطالعات گذشته یا نظر شخص متخصص موضوع موردمطالعه درباره میزان اهمیت اثرات شامل در $\boldsymbol{\theta}_2$ اطلاعاتی موجود باشد که دال بر یکسان نبودن شانس شامل شدن در $\boldsymbol{\theta}_2$ باشد، تابع‌های وزن دیگری از پارامترهای $\theta_p, \ldots, \theta_{m+1}$ را می‌توان در نظر گرفت. همچنین برای تعیین تابع وزن از اصل سلسله مراتبی اثرات می‌توان کمک گرفت. تابع وزن g را طوری می‌توان تعیین کرد که به اثرات مرتبه پایین و بالا وزن متفاوتی نسبت دهد. برای مثال در طرح 2^4 مثال ۲.۳، اثرات متقابل سه و چهار عاملی طبق اصل سلسله مراتبی از اهمیت کمتری نسبت به اثرات متقابل دو عاملی که در R_0 نیستند، برخوردارند. تعیین تابع وزن g به‌طوری که نقش اثرات متقابل سه و چهار عاملی را در تابع زیر انتگرال رابطه (۱.۳) کم رنگ کند، گزینه مناسبی می‌تواند باشد.

مثال ۳.۳. فرض می‌کنیم در مثال ۲.۳ اثرات متقابل مراتب سه و بالاتر قابل اغماض باشند. این فرض در مباحث غربالگری به‌وسیله‌ی طرح‌های کسری، طبق اصل سلسله مراتبی و اصل تنک بودن اثرات منطقی خواهد بود [۱۱]. همچنین فرض کنیم که طبق اطلاعات متخصص موضوع موردمطالعه، می‌دانیم که اثرات متقابل دو عاملی بین عامل F_4 و سایر عوامل وجود ندارد و اثر متقابل دو عاملی F_2F_3 به‌اشتباه قابل اغماض در نظر گرفته شده است (یعنی اثر متقابل F_2F_3 به‌اشتباه در مجموعه R_0 قرار داده شده است). ازاین‌رو داریم $\boldsymbol{\theta}_2 = (\theta_7, 0, \ldots, 0)$ که در آن θ_7 متناظر با اثر متقابل F_2F_3 است. در این حالت تابع وزن $g(\theta_7, \ldots, \theta_{15}) = |\theta_7|$ مناسب به نظر می‌رسد. این تابع وزن به θ_7 وزن معادل با قدر مطلق اثر را نسبت می‌دهد. با این انتخاب به‌سادگی می‌توان نشان داد که انتگرال رابطه (۱.۳) به‌صورت

$$L(\mathrm{Z}_1) = \sigma^{2(m+1)} det\big((\mathrm{Z}_1^T\mathrm{Z}_1)^{-1}\big)\big(\alpha + \frac{1}{2\sigma^2}a_{77}\alpha^2\big)$$

است که در آن a_{77} نشان‌دهنده عنصر قطری ماتریس A متناظر با اثر متقابل F_2F_3 است. برای $n = 8, \ldots 15$ وقتی که $\sigma = 1$ و $\alpha = 1$ است، طرح‌های WI – بهینه در جدول ۳ ارایه شده‌اند.

۴ بحث و نتیجه‌گیری

در یک طرح عاملی کامل تعداد کل آزمایش‌ها با افزایش عامل‌ها سریع افزایش می‌یابد و امکان اجرای همه‌ی آزمایش‌ها را در عمل با مشکل مواجه می‌کند. در این حالت استفاده از طرح‌های عاملی کسری رایج است. برای انتخاب طرح کسری معیارهای متعددی وجود دارند. در این مقاله براساس میانگین وزنی دترمینان ماتریس میانگین مربعات خطا معیاری معرفی شد. برای طرح‌های کسری دوسطحی نشان داده شد که با انتخاب تابع وزن یکنواخت، این طرح‌های بهینه با طرح D –بهینه معادل‌اند.

بدیهی است اگر به هر طریقی مثلا از مطالعات گذشته یا نظر شخص متخصص موضوع موردمطالعه درباره میزان اهمیت اثرات نادیده‌گرفته‌شده اطلاعاتی موجود باشد، تابع‌های وزن دیگری را می‌توان در نظر گرفت. این مساله انعطاف‌پذیر و کلی‌بودن معیار معرفی‌شده در این مقاله را نشان می‌دهد. ممکن است بتوان تابع‌های وزنی را پیدا کرد که طرح‌های بهینه معادل طرح‌های حاصل از معیارهایی مانند A–بهینه و غیره را نتیجه می‌دهد. در این‌صورت معیار معرفی‌شده در این مقاله را می‌توان تعمیم و حالتی کلی از بعضی از معیارهای موجود در نظر گرفت. به نظر نویسندگان این می‌تواند موضوع تحقیقی جالبی برای مطالعات آینده باشد.

برای یافتن طرح بهینه در یک طرح چهارعاملی دوسطحی برای تعداد اجراهای بین ۸ تا ۱۵ از الگوریتم جستجوی کامل استفاده شد. واضح است که با افزایش عامل‌ها استفاده از این الگوریتم به دلیل حجم زیاد محاسبات امکان پذیر نیست. در این حالت می‌توان از الگوریتم‌های جستجوی دنباله‌ای، تصادفی و نوردیدن استفاده کرد [۱۲]. نتایج الگوریتم جستجوی کامل بهینه‌ی دقیق است. اما نتایج سایر الگوریتم‌ها تقریبی است.

۵ تقدیر و تشکر

از داوران محترم برای پیشنهادهای ارزنده‌ای که موجب بهبود و ارائه بهتر مقاله شد و از هیئت تحریریه و سردبیر محترم مجله کمال تشکر را داریم.

فهرست منابع

[1] Box G. E. P. and Hunter J. S., *The ۲^k fractional factorial designs: part I*, Technometrics, 3 (1961a), 311-351.

[2] Box G. E. P. and Hunter J. S., *The ۲^k fractional factorial designs: part II*, Technometrics, 3 (1961b), 449-458.

[3] Cheng C. S., *Orthogonal arrays with variable numbers of symbols*, Ann. Stat., 8 (1980), 447-453.

[4] Fedorov V. V., *Theory of optimal experiments*, Academic Press, New York, (1972).

[5] Lehmann E. L. and Romano J., *Testing Statistical Hypotheses,* Springer, New York, (2005).

[6] Lin D. K. J. and Zhou J., *D-optimal minimax fractional factorial designs*, Can. J. Stat., 41 (2013), 325-340.

[7] Pukelsheim F., *Optimal design of experiments*, Wiley, New York, (1993).

[8] Rao C. R., *Linear Statistical Inference and its Applications,* Wiley, New York, (1973).

[9] Tang B. and Zhou J., *Existence and construction of two-level orthogonal arrays for estimating main effects and some specified two-factor interactions*, Stat. Sin., 19 (2009), 1193-1201.

[10] Wilmut M. and Zhou J., *D-optimal minimax design criterion for two-level fractional factorial designs*, J. Stat. Plan. Inference., 141 (2011), 576-587.

[11] Wu C. F. J. and . Hamada M. S, *Experiments: planning, analysis and optimization*, Wiley, New York, (2009).

[12] Yin Y. and Zhou J., *Minimax design criterion for fractional factorial designs*, Ann. Inst. Stat. Math. 67 (2015), 673-685.

A criterion for selecting a two level fractional factorial design

Nabaz Esmailzadeh [§] **, Shayda Ramezani**

Department of Statistics, University of Kurdistan, Sanandaj, Iran

Received: 2019/06/17 Accepted: 2020/10/24

Communicated by: F. Azarpanah

Abstract: Application of fractional factorial design is common in experiments with a large number of factors. Choosing the appropriate fraction is an important issue in the fractional designs literature. There are different criteria based on different perspectives. In this paper, a criterion based on minimization of a weighted function of the mean squared error matrix of the least squares estimators of a pre-specified model is introduced. In two-level fractional designs, the criterion for the uniform weight function is calculated and shown to be equivalent to the well-known D-optimal design. Finally, the method is described with examples.

Keywords: Fractional factorial design, Optimal design, Two level design.

Mathematics Subject Classification (2010): 62k05; 62k15.

[§]Corresponding author.
E-mail addresses: n.esmailzadeh@uok.ac.ir (N. Esmailzadeh), shaydaramezani67@gmail.com (S. Ramezani).