# بررسی کاربرد رگرسیون ریج در علوم پزشکی

نازنین فکری[۱]- حبیب‌الله اسماعیلی[۲]- حسن دوستی[۳]- رقیه پارسایی[۴]
[۱] گروه بهداشت عمومی، دانشکده علوم پزشکی نیشابور
[۲] گروه آمار زیستی و عضو مرکز تحقیقات علوم بهداشتی دانشکده بهداشت، دانشگاه علوم پزشکی مشهد
[۳] گروه آمار زیستی، دانشکده بهداشت، دانشگاه علوم پزشکی مشهد
[۴] گروه آمار زیستی و اپیدمیولوژی دانشکده بهداشت، دانشگاه علوم پزشکی اصفهان

**چکیده**

روش‌های رگرسیونی یکی از پر کاربردترین روش‌های آنالیز آماری در علوم پزشکی است. وجود همبستگی بین متغیرهای پیشگو باعث بوجود آمدن همخطی می شود. بنابراین برآورد پارامترها برمبنای کمترین مربعات باقیمانده نادرست است. در این مقاله سعی بر این است تا رگرسیون ریج، به عنوان یکی از کاراترین و موثرترین روش‌های مقابله با مشکل هم خطی معرفی گردد. در انتها روش فوق را با یک مثال کاربردی که روی داده‌های مربوط به عوامل خطر بیماری‌های قلبی انجام شده است، توضیح می‌دهیم.

**واژه‌های کلیدی:** هم خطی، رگرسیون ریج، کمترین مربعات باقیمانده.

## ۱ مقدمه

رگرسیون یکی از روش‌های متداول و قدیمی برای بررسی رابطه‌ی بین متغیرهای قابل اندازه‌گیری است. به ویژه رگرسیون خطی که تقریباً در همه‌ی شاخه‌های علوم مانند: علوم اجتماعی اقتصادی، علوم فیزیکی و زیست شناسی، پزشکی، تجارت، تکنولوژی و غیره کاربرد فراوان دارد. در مدل‌های رگرسیون سعی بر این است که از طریق یک یا چند متغیر پیشگو، متغیر پاسخ پیش‌بینی شود[۱]. مدل استاندارد برای رگرسیون خطی چندگانه به شکل

$$\underline{Y} = X\underline{\beta} + \underline{e} \qquad (۱)$$

است که در آن $\underline{Y}$ بردار $n \times ۱$ از متغیر پاسخ، $X$ یک ماتریس $n \times p$ از متغیرهای پیشگو با رتبه‌ی $p$، $\underline{\beta}$ بردار

$p \times ۱$ ضرایب و مجهول هستند. $\underline{e}$ نیز بردار $n \times ۱$ باقی‌مانده‌های مدل با میانگین صفر و واریانس ثابت $\sigma^۲$ است [۳].

با کمک داده‌های جمع آوری شده و به کارگیری روش کمترین مربعات باقی‌مانده[۱] (روش گاوس مارکوف)، ضرایب مجهول $\beta_۱, \cdots, \beta_p$ را می‌توان برآورد کرد و بدین ترتیب رابطه‌ی (۱) مشخص می‌شود. به عبارتی اگر تعداد مشاهدات $n$ باشد آنگاه داریم[۱]:

$$y_i = b_۱ x_{i۱} + b_۲ x_{i۲} + \cdots + b_p x_{ip} + e_i,$$

که $i = ۱, ۲, \ldots, n$. و یا به صورت ماتریس:

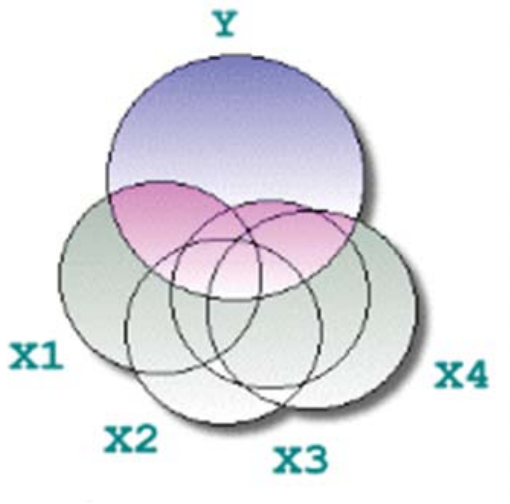
$$\underline{Y} = X\underline{b} + \underline{e}.$$

ضرایب $b$ در این حالت برآورد ضرایب $\beta$ و نااریب با کمترین واریانس هستند.

درنظر نگرفتن پیش‌فرض‌ها در رگرسیون خطی می‌تواند روی نتیجه‌گیری‌ها اثر نامطلوب گذاشته و نتایجی نادرست بدست آید. یکی از این پذیره‌ها وجود همخطی در ماتریس داده‌ها است.

## ۲ هم خطی چیست؟

$p$ متغیر را زمانی هم‌خط گویند که حداقل یکی از بردارهای ماتریس $X$ یک ترکیب خطی از بردارهای دیگر باشد[۴]. به عبارتی دیگر، هرگاه دو متغیر با یکدیگر هم خطی داشته باشند، یعنی یکی از آنها، در معادله ی رگرسیون، حرف جدیدی برای گفتن ندارد و حضورش در معادله توجیه نمی‌شود. شکل۱، نمونه ای از وجود هم‌خطی بین متغیرهای توضیحی را نشان می‌دهد. ستون های ماتریس $X$ به صورت $x_۱, \ldots, x_p$ هستند به طوری که هر یک از




شکل ۱: وجود هم‌خطی بین متغیرهای توضیحی $X_۱$ تا $X_۴$

این بردارها مقادیر مشاهده شده برای هر کدام از متغیرهای پیشگو هستند. اگر به ازای ثابت‌های $a_۱, \ldots, a_p$ یک رابطه‌ی خطی کامل، یعنی $\sum_{i=۱}^{p} a_i x_i = ۰$ وجود داشته باشد، آنگاه می‌گوییم «هم‌خطی کامل» وجود دارد و اگر این رابطه‌ی خطی تقریبی باشد، یعنی $\sum_{i=۱}^{p} a_i x_i \approx ۰$ آنگاه گوییم «هم‌خطی ناقص» وجود دارد. در صورت وجود هم‌خطی کامل، معکوس ماتریس $X^T X$ وجود نخواهد داشت و نیز کوچک‌ترین مقدار ویژه‌ی ماتریس $X^T X$ برابر صفر است. هم‌خطی ناقص نیز بدین معنی است که $X^T X$ تقریباً منفرد است. و کوچک‌ترین مقدار ویژه‌ی ماتریس $X^T X$ نزدیک صفر می‌باشد.

## ۳ آثار هم خطی

برخی از مهمترن نشانه‌های حضور هم خطی بین متغیرهای پیشگو به صورت زیر است:

- با حذف یا اضافه کردن یک متغیر پیشگو، یا با

[۱] Least square residuals

ورود و خروج یک مشاهده به مدل، تغییرات زیادی در ضرایب برآورد شده‌ی رگرسیون بوجود می‌آید.

- نتایج غیر معمول در آزمون معنی داری ضرایب رگرسیون، دیده می‌شود.
- ضرایب رگرسیون برآورد شده، علامتی مخالف آنچه مورد انتظار است دارند.
- فواصل اطمینان پرعرض برای ضرایب رگرسیونی مربوط به متغیرهای توضیحی مهم موجود در مدل، برآورد می‌شود.
- انحراف استاندارد برآورد شده‌ی ضرایب رگرسیون، زیاد می‌شود.

موارد ذکر شده، حضور هم خطی را به صورت کمی اندازه‌گیری نمی‌کنند [۴].

## ۴ روش‌های تشخیص هم‌خطی

از جمله روش های تشخیص هم خطی می توان به موارد زیر اشاره کرد:

۱. رسم نمودار leverage یا نمودار پراکندگی

۲. استفاده از آزمون‌های نسبت-F[۲] بالا و یا نسبت-T[۳] کوچک

۳. استفاده از شاخص عامل تورم واریانس[۴] (VIF). این شاخص برای $k = ۱, \ldots, p$ برابر است با:

$$VIF_k = (۱ - R_k^۲)^{-۱},$$

که در آن $R_k^۲$ مقدار ضریب تعیین چندگانه بین متغیر ها است، زمانی که $X_k$ را روی سایر متغیرها برگشت دهیم [۲]. با توجه به رابطه‌ی بالا چنان چه $R_k^۲ = ۰$ باشد، آنگاه مقدار $VIF_k = ۱$ است. بنابراین اگر $X_k$ رابطه‌ی خطی با متغیرهای پیشگوی دیگر نداشته باشد، مقدار $VIF_k = ۱$ خواهد بود.

## ۵ روش‌های مقابله با هم‌خطی

برای حذف یا کم کردن هم خطی بین متغیر های پیشگو روش های مختلفی ارائه شده است که به طور کلی این روش‌ها شامل حذف متغیر پیشگویی که باعث ایجاد هم خطی شده، جمع آوری بیشتر و مجدد داده و یا استفاده از روش‌هایی غیر از کمترین مربعات باقیمانده است. استفاده از روش‌های غیر از کمترین مربعات باقیمانده از نظر کاربردی و دقت، دارای کارایی بیشتری هستند و از کاهش توان آزمون جلوگیری می‌کنند. یکی از این روش‌ها برای مقابله با اثرات نامطلوب هم‌خطی استفاده از برآوردگرهای اریب است که رگرسیون ریج، یکی از مهم‌ترین و کاراترین این روش‌هاست.

## ۶ رگرسیون ریج

هارل و کنارد [۳] اولین بار در مقاله ای در مجله‌ی مهندسی شیمی، بیان کردند زمانی که در مدل هم‌خطی وجود داشته باشد، برای برآورد $\beta$، می‌توان به جای برآوردگرهای کمترین مربعات باقی‌مانده از برآوردگر زیر استفاده کرد:

$$\widehat{\beta^*} = [X'X + KI]^{-۱} X'Y,$$

[۲] F-ratio
[۳] T-ratio
[۴] Variance Inflation Factor

برآوردهایی که با $K \geq 0$ بدست می‌آیند، شباهت‌های ریاضی زیادی با تعریف توابع پاسخ درجه دوم دارند. به این دلیل، برآوردها و تحلیل‌هایی که با $\widehat{\beta^*}$ انجام می‌شود " رگرسیون ریج " نامیده می‌شوند[۳].
رابطه‌ی برآورد ریج با برآورد کمترین مربعات باقی‌مانده، به شکل زیر است:

$$\tilde{b} = \widehat{\beta^*} = Z\hat{\beta}.$$

به عبارتی برآوردگر ریج یک ترکیب خطی از برآوردگر کمترین مربعات باقی‌مانده است، به طوری که

$$Z = [X'X + KI]^{-1}X'X.$$

در روش کمترین مربعات باقی‌مانده، فرض کردیم $\hat{b}$ برآوردی نااریب برای $\beta$ است و خاصیت گاوس مارکوف، مینیمم بودن واریانس این برآوردگر در کلاس برآوردگرهای خطی نااریب را بیان می‌کند.
در محاسبه‌ی برآوردگر ریج از فرض نااریب بودن چشم‌پوشی شده و برآوردگر اریبی برای $\beta$ بدست می آید که دارای واریانس کمتری نسبت به براوردگر کمترین مربعات باقی‌مانده است. بنابراین مقدار برآورد در این حالت پایدارتر خواهد بود[۴]، یعنی با حذف یا افزودن یک متغیر پیشگوی جدید به مدل و یا... ضرایب برآوردشده‌ی مدل تغییر چندانی نخواهند کرد.

در محاسبه‌ی برآوردگر ریج ($\tilde{b}$)، با فرض این که ستون‌های ماتریس $X$ استاندارد شده باشند، از حل معادلات نرمال داریم:

$$\tilde{b}(k) = (x'x + kI)^{-1}x'y,$$

که در آن $k$ مقداری مثبت است که بایستی توسط تحلیل‌گر با توجه به یک سری معیارها تعیین شود.
با توجه به این نکته که $MSE(\underline{\tilde{b}}) = \text{variance} + (\text{bias})^2$. بنابراین برای انتخاب $k$ مقداری را برمی‌گزینیم که کاهش در واریانس برآوردگر اریب، بیش از افزایش مربع اریبی باشد. در نتیجه MSE آن کمتر از واریانس برآوردگر نااریب خواهد بود[۱].

## ۷ کاربرد

یک مطالعه مقطعی، برای بررسی ارتباط بین شاخص‌های چاقی و تیتر آنتی بادی شوک حرارتی (HSP-27) در دانشگاه علوم پزشکی مشهد روی سه گروه متشکل از ۵۰ نفر با وزن طبیعی، ۵۰ نفر دارای اضافه وزن و ۱۰۰ نفر که خیلی چاق هستند، انجام شده است، متغیرهای اندازه‌گیری شده در این مطالعه فشار خون، سن، جنس، چربی خون، قند خون و تعدادی از عوامل خطر بیماری‌های قلبی-عروقی بوده‌اند [۴]. در مطالعه فوق وجود همخطی بین متغیرها در نظر گرفته نشده است. بین متغیرها در گروه اول(افراد با وزن طبیعی)، VIF بررسی و همخطی شدید مشاهده شد. با استاندارد کردن متغیرهای پیشگو سعی در از بین بردن هم‌خطی (در همان گروه اول) نمودیم که نتایج در جدول ۱ آمده است. با توجه به جدول۱ و انجام رگرسیون روی داده های استاندارد شده مشخص شد که هنوز همخطی وجود داشته و در رگرسیون معمولی تنها متغیر فشار خون بر HSP-27، معنی‌دار شد.

بنابراین جهت رفع هم‌خطی، از رگرسیون ریج استفاده نمودیم. با توجه به مطالبی که در قسمت رگرسیون ریج آمده، ضریب $K$ محاسبه و با استفاده از فرمول $\tilde{b}(k) = (x'x + kI)^{-1}x'y$ پارامترهای مدل مجدداً برآورد شد.

جدول ۲ به‌ازای مقادیر مختلف $k$ برآورد ضرایب

رگرسیون و مقدار VIF را نشان می‌دهد. همانطور که ملاحظه می شود، مقدار VIF برای $k = ۰/۱۶$ به میزان یک و یا کمتر رسیده که بیانگر عدم وجود همخطی است. از طرفی، ضرایب برآورد شده، در حالت $k = ۰/۱۶$ به مقدار پایدارتری رسیده و تغییرات آن‌ها کمتر می‌باشد.

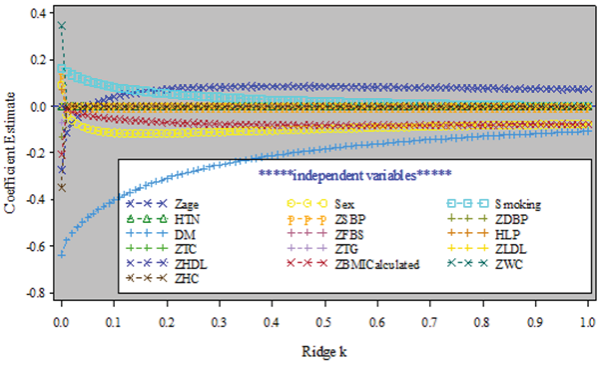
جدول ۱: برآورد ضرایب مدل رگرسیون معمولی در تاثیر متغیر های استاندارد شده بر HSP-27

| $Variable$ | $DF$ | $Parameter\ Estimate$ | $Standard\ Error$ | $tValue$ | $Pr > \|t\|$ | $Variance\ Inflation$ |
|---|---|---|---|---|---|---|
| $intercept$ | ۱ | ۰/۰۹۸۳۴ | ۱/۲۸۱۵۱ | ۰/۰۸ | ۰۹۳۹۳ | ۰ |
| $Zage$ | ۱ | −۰/۲۷۲۱۳ | ۰/۱۸۹۶۴ | −۱/۴۳ | ۰/۱۶۰۷ | ۳/۷۶۵۲۱ |
| $Sex$ | ۱ | ۰/۰۹۲۲۳ | ۰/۲۷۲۴۵ | ۰/۳۴ | ۰/۷۳۷۰ | ۱/۷۷۲۴۷ |
| $Smoking$ | ۱ | ۰/۱۶۰۸۳ | ۰/۳۴۵۴۶ | ۰/۴۷ | ۰/۶۴۴۶ | ۱/۵۳۴۴۵ |
| $HTN$ | ۱ | ۰/۰۰۲۶۸ | ۰/۰۰۱۲۲ | ۲/۲۰ | ۰/۰۳۴۷ | ۱۶/۹۵۹۵۶ |
| $ZSBP$ | ۱ | ۰/۱۲۸۲۳ | ۰/۱۷۹۷۸ | ۰/۷۱ | ۰/۴۸۰۷ | ۴۵۵۳۶۰ |
| $ZDBP$ | ۱ | −۰/۱۳۰۱۹ | ۰/۱۷۹۹۸ | −۰/۷۲ | ۰/۴۷۴۵ | ۴۵۶۶۰۴ |
| $DM$ | ۱ | −۰/۶۳۶۴۸ | ۰/۴۶۵۴۷ | −۱/۳۷ | ۰/۱۸۰۸ | ۱/۵۲۵۵۷ |
| $ZFBS$ | ۱ | ۰/۰۰۱۰۸ | ۰/۰۰۰۸۰۶۱۲ | ۱/۳۴ | ۰/۱۸۸۲ | ۶/۵۵۱۷۲ |
| $HLP$ | ۱ | ۰/۰۶۹۳۹ | ۰/۱۲۶۶۹ | ۰/۵۵ | ۰/۵۸۷۶ | ۱۳۷۴۷۵ |
| $ZTC$ | ۱ | −۰/۰۰۰۴۵۴۷۷ | ۰/۰۰۰۳۰۴۲۸ | −۱/۴۹ | ۰/۱۴۴۵ | ۲/۰۸۱۰۵ |
| $ZTG$ | ۱ | −۰/۰۷۰۵۴ | ۰/۱۲۶۲۸ | −۰/۵۶ | ۰/۵۸۰۲ | ۱۳۷۰۹۰ |
| $ZLDL$ | ۱ | ۰/۰۰۰۳۹۴۶۶ | ۰/۰۰۰۴۰۳۰۹ | ۰/۹۸ | ۰/۳۳۴۷ | ۲/۴۷۹۶۷ |
| $ZHDL$ | ۱ | −۰/۰۰۰۷۶۶۷۳ | ۰/۰۰۰۷۰۰۱۶ | −۱/۰۹ | ۰/۲۸۱۶ | ۵/۶۳۰۶۲ |
| $Calculated\ ZBMI$ | ۱ | −۰/۲۰۴۵۷ | ۰/۳۷۰۹۹ | −۰/۵۵ | ۰/۵۸۵۱ | ۱/۸۶۷۹۵ |
| $ZWC$ | ۱ | ۰/۳۴۶۵۳ | ۰/۲۴۲۴۹ | ۱/۴۳ | ۰/۱۶۲۴ | ۷۵۵۹۳۳ |
| $ZHC$ | ۱ | −۰/۳۴۷۰۱ | ۰/۲۴۲۴۹ | −۱/۴۳ | ۰/۱۶۱۸ | ۷۵۵۹۹۸ |

جدول ۲: برآورد ضرایب رگرسیونی به ازای مقادیر مختلف $k$.

| obs | type | k | $b_\circ$ | Zage | sex | smoking | HTN | ZSBP | ZDBP | DM | ZFBS | HLP | ZTC | ZTG |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ۲۲ | VIF | ۰/۱۰ | | ۱/۰۰۲۴۸ | ۱/۰۰۵۰۵ | ۰/۹۷۷۵۹ | ۱/۰۵۰۷۷ | ۰/۴۲ | ۰/۴۱ | ۰/۹۹۸۴۲ | ۵۰۶۲۸/۱ | ۰/۶۹ | ۱/۱۷۳۶۱ | ۰/۶۹ |
| ۲۳ | $\tilde{b}(k)$ | ۰/۱۰ | ۰/۳۱۴۲۳ | ۰/۰۴۱۶۱ | −۰/۱۱۳۶۹ | ۰/۰۸۲۲۳ | ۰/۰۰۰۹ | −۰/۰۰ | −۰/۰۰ | −۰/۴۰۲۴۷ | ۰/۰۰۰۶۷ | −۰/۰۰ | −۰/۰۰۰۲۱ | −۰/۰۰ |
| ۲۴ | $VIF$ | ۰/۱۱ | | ۰۹۵۸۰۰ | ۰/۹۷۶۳۹ | ۰/۹۵۰۰۲ | ۱/۳۳۵۸ | ۰/۳۸ | ۰/۳۸ | ۰/۹۶۹۲۷ | ۱/۶۸۳۰۳ | ۰/۶۳ | ۱/۱۳۳۳۸ | ۰/۶۳ |
| ۲۵ | $\tilde{b}(k)$ | ۰/۱۱ | ۰/۲۹۴۹۷ | ۰/۰۴۷۱۰ | −۰/۰۱۱۴۹۸ | ۰/۰۷۸۳۷ | ۰/۰۰۰۸ | −۰/۰۰ | −۰/۰۰ | −۰/۳۹۰۵۶ | ۰/۰۰۰۶۴ | −۰/۰۰ | −۰/۰۰۰۲۰ | −۰/۰۰ |
| ۲۶ | $VIF$ | ۰/۱۲ | | ۰۹۱۸۸۱ | ۰/۹۴۹۳۸ | ۰/۹۲۴۰۳ | ۱/۱۹۴۲ | ۰/۳۶ | ۰/۳۵ | ۰/۹۴۱۷۸ | ۱/۵۵۷۶۶ | ۰/۵۸ | ۱/۰۹۵۶۷ | ۰/۵۸ |
| ۲۷ | $\tilde{b}(k)$ | ۰/۱۲ | ۰/۲۷۶۵۹ | ۰/۰۵۱۸۲ | −۰/۱۱۵۸۸ | ۰/۰۷۴۸۳ | ۰/۰۰۰۸ | −۰/۰۰ | −۰/۰۰ | −۰/۳۷۹۴۰ | ۰/۰۰۰۶۱ | −۰/۰۰ | −۰/۰۰۰۲۰ | −۰/۰۰ |
| ۲۸ | $VIF$ | ۰/۱۳ | | ۰/۸۸۳۸۷ | ۰/۹۲۳۸۳ | ۰/۸۹۹۴۶ | ۱/۰۷۶۰ | ۰/۳۳ | ۰/۳۳ | ۰/۹۱۵۷۹ | ۱/۴۴۷۰۵ | ۰/۵۴ | ۱/۰۶۰۲۱ | ۰/۵۴ |
| ۲۹ | $\tilde{b}(k)$ | ۰/۱۳ | ۰/۲۵۹۰۲ | ۰/۰۵۵۹۱ | −۰/۱۱۶۴۷ | ۰/۰۷۱۵۶ | ۰/۰۰۰۷ | −۰/۰۰ | −۰/۰۰ | −۰/۳۶۸۹۱ | ۰/۰۰۰۵۸ | −۰/۰۰ | −۰/۰۰۰۲۰ | −۰/۰۰ |
| ۳۰ | $VIF$ | ۰/۱۴ | | ۰/۸۵۲۳۸ | ۰/۸۹۹۵۹ | ۰/۸۷۶۱۵ | ۰/۹۷۶۳ | ۰/۳۱ | ۰/۳۱ | ۰/۸۹۱۱۷ | ۱/۳۴۸۸۸ | ۰/۵۰ | ۱/۰۲۶۷۸ | ۰/۵۰ |
| ۳۱ | $\tilde{b}(k)$ | ۰/۱۴ | ۰/۲۴۲۲۱ | ۰/۰۵۹۴۶ | −۰/۱۱۶۸۱ | ۰/۰۶۸۵۳ | ۰/۰۰۰۷ | −۰/۰۰ | −۰/۰۰ | −۰/۳۵۹۰۲ | ۰/۰۰۰۵۵ | −۰/۰۰ | −۰/۰۰۰۲۰ | −۰/۰۰ |
| ۳۲ | $VIF$ | ۰/۱۵ | | ۰/۸۲۳۷۷ | ۰/۸۷۶۵۳ | ۰/۸۵۴۰۱ | ۰/۸۹۱۳ | ۰/۲۹ | ۰/۲۹ | ۰/۸۶۷۸۰ | ۱/۲۶۱۲۸ | ۰/۴۶ | ۰/۹۹۵۱۹ | ۰/۴۷ |
| ۳۳ | $\tilde{b}(k)$ | ۰/۱۵ | ۰/۲۲۶۱۳ | ۰/۰۶۲۵۶ | −۰/۱۱۶۹۵ | ۰/۰۶۵۷۱ | ۰/۰۰۰۷ | −۰/۰۰ | −۰/۰۰ | −۰/۳۴۹۶۸ | ۰/۰۰۰۵۳ | −۰/۰۰ | −۰/۰۰۰۱۹ | −۰/۰۰ |
| ۳۴ | $VIF$ | ۰/۱۶ | | ۰/۷۹۷۵۷ | ۰/۸۵۴۵۶ | ۰/۸۳۲۹۳ | ۰/۸۱۸۴ | ۰/۲۸ | ۰/۲۸ | ۰/۸۴۵۵۷ | ۱/۱۸۲۷۳ | ۰/۴۳ | ۰/۹۶۵۲۷ | ۰/۴۴ |
| ۳۵ | $\tilde{b}(k)$ | ۰/۱۶ | ۰/۲۱۰۷۳ | ۰/۰۶۵۲۷ | −۰/۱۱۶۹۳ | ۰/۰۶۳۰۷ | ۰/۰۰۰۶ | −۰/۰۰ | −۰/۰۰ | −۰/۳۴۰۸۳ | ۰/۰۰۰۵۰ | −۰/۰۰ | −۰/۰۰۰۱۸ | −۰/۰۰ |

همچنین نمودار اثر ریج[۵] نیز برآورد ضرایب هر یک از متغیرها را به ازای مقادیر مختلف $k$ نشان می‌دهد. این نمودار شکل دو بعدی است که با توجه به جدول ۲ رسم می‌شود و در آن، با افزایش مقدار $k$ برآورد ضرایب پایدارتر می‌شود [۲].




شکل ۲: نمودار اثر ریج

## ۸ بحث و نتیجه‌گیری

در صورت وقوع هم خطی، بهتر است از روش رگرسیون ریج برای برآورد ضرایب مدل استفاده شود. که در این حالت، MSE، حتی کمتر از ضرایب در روش کمترین مربعات باقی‌مانده است و دقت مدل بیشتر می‌شود.

## تقدیر و تشکر

نویسندگان از تمام کارکنان کتابخانه‌ی دانشکده‌ی بهداشت، دانشگاه علوم پزشکی مشهد و نیز کارکنان کتابخانه‌ی استاد فاطمی، دانشکده‌ی علوم ریاضی، دانشگاه فردوسی مشهد تقدیر و تشکر می‌نمایند.

[۵] Ridge trace

## مراجع

[۱] آذرنوش، ح. ع. و دوج، ی. (۱۳۷۹)، روش های جایگزین در رگرسیون. ترجمه. دانشگاه فردوسی مشهد.

[2] Hoerl, A. E and Kennard, R. W. (1970a). Ridge Regression: Applications to Nonorthogonal Problems. Technometrics: 69-82.

[3] Hoerl, A. E. and Kennard, R. W. (1970b). Ridge Regression: Biased Estimation for Nonorthogonal Problems. Technometrics: 55-67.

[4] Kutner, M., Nachtsheim, C., Neter, J. and Li, W. (2005). Applied Linear Statistical Models. 5thed. New York: McGraw-Hill.

[5] Tavallaie,S., Rahsepar, A. A. et al. (2012). Association between indices of body mass and antibody titers to heat-shock protein-27 in healthy subjects. Clinical Biochemistry, Vol. 45, 1-2, 144-147.