
بسم الله الرحمن الرحیم

دانشکده علوم ریاضی
گروه آمار

رساله
جهت اخذ درجه دکتری آمار

عنوان
برآورد در مدل های خطی جزئی

تهیه و تنظیم
مهدی روزبه

استاد راهنما
دکتر حسینعلی نیرومند

استاد مشاور
دکتر محمد آرشی

تابستان ۱۳۹۰

بسمه تعالی

مشخصات رساله دانشجویان دکتری

دانشگاه فردوسی مشهد

عنوان پایان نامه : برآورد در مدل های خطی جزئی	
نویسنده: مهدی روزبه	**استاد راهنما: دکتر حسینعلی نیرومند**
دانشکده : علوم ریاضی	**گروه: آمار**
رشته تحصیلی : آمار ریاضی	**تاریخ دفاع: ۹۰/۶/۳۰**
مقطع تحصیلی : کارشناسی ○ کارشناسی ارشد ○ دکتری ●	**تعداد صفحات : ۱۵۰ صفحه**

چکیده :

در این رساله، از روش تفاضلی برای حذف اثر تابع ناپارامتری در مدل های خطی جزئی به منظور برآورد قسمت خطی استفاده می کنیم. سپس، بعد از برآورد قسمت خطی، می توان از یکی از تکنیک های ناپارامتری برای برآورد قسمت غیر خطی استفاده نمود.

ایده تفاضلی کردن برای حذف اثر تابع ناپارامتری در رگرسیون ناپارامتری چندان جدید نیست. این ایده در مدل های خطی جزئی توسط اهن و پاول در سال ۱۹۹۴ بکار برده شد. در حالیکه روش تفاضلی در مدل رگرسیون ناپارامتری محض قدیمی تر بوده و اولین بار توسط رایس در سال ۱۹۸۴ بمنظور برآورد واریانس جملات خطا بکار برده شد. همچنین ثابت می شود که برآوردگرهای تفاضلی بهینه هستند، در واقع، ونگ و همکارانش (۲۰۰۷) ثابت کردند که با استفاده از روش تفاضلی، برآوردگر قسمت خطی بطور مجانبی کارا بوده و برآوردگر غیر خطی دارای نرخ همگرایی مینیماکس خواهد بود.

در ادامه یک روش برآوردیابی اریب را در مدل های خطی جزئی وقتی که ماتریس **X'X** دچار بد شرطی است، بکار می بریم. در واقع، مفهوم رگرسیون ریج را که در سال ۱۹۷۰ توسط هرل و کنارد برای غلبه بر مشکل همخطی در مسائل رگرسیونی پیشنهاد شد، مورد استفاده قرار می دهیم.

واژه های کلیدی : برآوردگر تفاضلی، برآوردگر ریج، برآوردگر شدنی، برآوردگر کمترین مربعات تعمیم یافته، تابع زیان متعادل، همخطی، محدودیت های خطی، مدل های خطی جزئی، مدل های نیمه پارامتری به ظاهر ناهمبسته، هموارساز کرنل	**امضای استاد راهنما:** **تاریخ**

In the name of God

Thesis details
Ferdowsi university of Mashhad

Title : **Estimation of Partial Linear Models.**

Student name : Mahdi Roozbeh	**Supervisor : Dr. H. A. Niroumand**
School : Mathematical of Sciences	**Department: Statistics**
Education Field : Mathematical Statistics	**Date of defense: 22 Sep., 2011**
For Degree: BS ○ **MSc** ○ **Ph.D** ●	**Number of page : 150**

Abstract:

In this thesis, we use differencing method to remove the nonparametric part of partial linear model (PLM) in order to estimate the linear part. After estimation of the linear part, a variety of nonparametric techniques can be applied to estimate the non-parametric part.

The idea of differencing to remove the nonparametric effect in nonparametric regression models is not new. In the PLM, it has been examined by Ahn and Powell (1993). In a pure nonparametric regression setting, the idea of differencing has a longer history have been used by Rice (1984) to obtain estimators of the residual variance. The difference-based estimation procedure is optimal in the sense that the estimator of the linear component is asymptotically efficient and the estimator of the nonparametric component is asymptotically minimax rate optimal for the PLM (Wang *et al*. 2007).

We will examine a biased estimation technique to follow when the matrix $\mathbf{X}'\mathbf{X}$ appears to be ill-conditioned in the PLM. For the purpose of this thesis, we employ the ridge regression concept proposed in the 1970's by Hoerl and Kennard to combat the multicollinearity in regression problems.

Supervisor's Signature **Date**	**Keywords :** Balanced loss function; Differencing estimator; Feasible estimator; Generalized restricted least squares estimator; Kernel smoothing; Linear restrictions; Multicollinearity; Partial linear model; Ridge estimator; Seemingly unrelated Semiparametric model

....از بیابان بوی گندم مانده است

عشق روی دست مردم مانده است

آسمان بازیچه ی طوفان ماست

....ابر نعش آه سرگردان ماست

باز هم یک روز طوفان می شود

هر چه می خواهد خدا آن می شود

می روم افتان و خیزان تا غدیر

باده ها می نوشم از جوشن کبیر

آب زمزم در دل صحرا خوش است

باده نوشی از کف مولا خوش است

فاش می گویم که مولایم علیست

آفتاب صبح فردایم علیست

هر که در عشق علی گم می شود

مثل گل محبوب مردم می شود

تا علی گفتم زبان آتش گرفت

پیش چشمم آسمان آتش گرفت

آسمان رقصید و بارانی شدیم

موج زد دریا و طوفانی شدیم

یا علی

بغض چندین ساله ی ما باز شد

یا علی گفتیم و عشق آغاز شد

یا علی گفتیم و دریا خنده کرد

عشق ما را باز هم شرمنده کرد

یا علی گفتیم و گلها وا شدند

عشق آمد قطره ها دریا شدند

یاعلی گفتیم و طوفانی شدیم

مست از آن دستی که می دانی شدیم

یاعلی گفتیم و طوفان جان گرفت

کوفه در تزویر خود پایان گرفت

کوفه یعنی دستهای ناتنی

کوفه یعنی مردهای منحنی

کوفه یعنی مرد آری مرد نیست

یا اگر هم هست صاحب درد نیست

عده ای رندان بازاری شدند

عده ای رسوایی جاری شدند

آن همه دستی که در شب طی شدند

........ابن ملجم های پی در پی شدند

از سکوت و گریه سرشارم علی

تا همیشه دوستت دارم علی

تقدیم خالصانه به پدرم

وزین ترین کلمه در دیباچه هستی، تقدیم به او که هر چه دارم مدیون زحمات بی دریغش است. کاش می توانستم ذره ای از گذشت و فداکاری اش را جبران نمایم.

تقدیم عاشقانه به مادرم

شگفت ترین زاده عشق الهی، گوهری بی همتا در بحر بی کران فداکاری ها، کسی که دامان پرمهرش پناهگاه دوران کودکی ام و قلب پاکش سرچشمه دعای خیر در زندگی ام بوده است.

فهرست مقالات مستخرج از متن پایان نامه

[۱] Roozbeh, M., Arashi, M and Niroumand, H. A. (2010). On restricted semi-parametric models. *Proceedings of the 10th Iranian Statistical Conference*, University of Tabriz, 365-376.

[۲] Roozbeh, M., Arashi, M. and Niroumand, H. A. (2010). Semiparametric ridge regression approach in partially Linear Models. *Communications in Statistics-Simulation and Computation*, **39**, 449-460.

[۳] Roozbeh, M., Arashi, M and Niroumand, H. A. (2011). Ridge regression methodology in partial linear models with Correlated errors. *Statistical Computation and Simulation*, **81**, 517–528

[۴] Roozbeh, M., Arashi, M and Gasparini, M. (2011). Seemingly unrelated ridge regression in Semiparametric Models. *Communications in Statistics-Theory and Methods*, in press.

[۶] Roozbeh, M., Arashi, M and Niroumand, H. A. (2011). On seemingly unrelated semiparametric models. *Proceedings of the 58th ISI World Statistics Congress*, Dublin, Ireland.

[۷] Roozbeh, M. (2011). Feasible ridge estimator in partially linear models. Jan Tinbergen Awards 2011, Trinity college, University of Dublin, Ireland.

پیشگفتار

مدل های خطی جزئی در دو دهه اخیر کاربردهای زیادی پیدا نموده و توجه بسیاری از محققان را به خود جلب کرده است. این مدل ها کاربرد زیادی در مسائل اقتصادی دارند که از مهمترین آن ها می توان به مدل تابع درآمد سرمایه انسانی (ویلیس،[۱] ۱۹۸۶) و منحنی دستمزد (بلنچفلاور و اسوالد،[۲] ۱۹۹۴) اشاره کرد. در هر دو مورد، لگاریتم درآمد شخصی به ویژگی های شخصیتی (مانند جنس و وضعیت تاهل) و اندازه سرمایه های انسانی فرد (مانند تحصیل و تجربه کار در بازار) وابسته است. در نظریه اقتصاد رابطه غیر خطی بین لگاریتم درآمد شخصی و تجربه کار در بازار مورد بررسی قرار می گیرد.

مدل های خطی جزئی نخستین بار توسط انگل[۳] و همکارانش در سال ۱۹۸۶ بکار برده شد. هدف آن ها بررسی رابطه بین مصرف برق ماهیانه Y با متغیرهای قیمت برق ماهیانه $X_۱$، درآمد ماهیانه $X_۲$ و دمای هوا t بود. آن ها حدس زدند که متغیر وابسته یعنی Y دارای یك رابطه خطی با متغیرهای $X_۱$ و $X_۲$ و دارای یك رابطه غیر خطی با متغیر t است. اشمالنسی و استاکر[۴] در سال ۱۹۹۹ از این مدل ها برای بررسی رابطه بین مصرف خانگی گازوئیل با متغیرهایی مانند درآمد، سن تعداد افرادی که در هر خانواده رانندگی می کنند، تعداد افراد خانواده، محال اقامت و نوع خانه در ایالت متحده استفاده نمودند. آن ها دریافتند که لگاریتم مصرف گازوئیل بر حسب گالن دارای یك رابطه غیر خطی با لگاریتم متغیرهای درآمد و سن، و دارای یك رابطه خطی با لگاریتم تعداد افرادی که رانندگی می کنند و تعداد افراد هر خانواده و یك رابطه خطی با سایر متغیرهاست.

تا کنون روش های متعددی جهت برآورد مدل های خطی جزئی ارائه شده است. بعنوان مثال، انگل و همکارانش (۱۹۸۶) و چن و شیائو[۵] (۱۹۹۱) از روش کمترین مربعات

---

[۱] Willis

[۲] Blanchflower and Oswald

[۳] Engle

[۴] Schmalensee and Stoker

[۵] Chen and Shiau

جریمه ای[۶]، کیوزیك[۷] (۱۹۹۲) با استفاده از روش مانده های جزئی[۸]، سورینی و وانگ[۹] (۱۹۹۲) و کارول[۱۰] و همکارانش (۱۹۹۷) با استفاده از روش درستنمایی نیمرخ[۱۱]، پارامترهای این مدل را برآورد نمودند. پس از آن، فن[۱۲] و همکارانش (۱۹۹۸) با فرض معلوم بودن قسمت خطی به برآورد تابع ناپارامتری پرداخته و سپس قسمت خطی را برآورد نمودند و برخی خواص این نوع برآوردگرها توسط مامن و همکارانش[۱۳] (۱۹۹۹) استخراج گردید. در مورد روش های پیشنهاد شده برای حل قسمت غیر خطی می توان به تکنیك برآورد به روش انتگرال حاشیه ای جستیم و اوستد[۱۴] (۱۹۹۴) و لینتن و نیلسن[۱۵] (۱۹۹۵)، الگوریتم پس برازش بوجا[۱۶] توسط هستی و تیبشیرانی[۱۷] (۱۹۸۹) و رویکرد تقریب بوسیله سری فوریه از آماتو[۱۸] و همکارانش (۲۰۰۲) اشاره نمود.

تمرکز این رساله بر برآورد پارامترهای مدل های خطی جزئی به روش تفاضلی[۱۹] است. روش تفاضلی مرتبه اول نخستین بار توسط رایس[۲۰] (۱۹۸۴) برای برآورد واریانس خطا در مدل رگرسیون ناپارامتری پیشنهاد شد. سپس هال[۲۱] و همکارانش (۱۹۹۰)، روش پیشنهادی توسط رایس در برآورد واریانس خطا را تعمیم داده و نشان دادندکه با بالا بردن مرتبه

---

[۶] Penalized least square method

[۷] Cuzick

[۸] Partial residual method

[۹] Severini and Wong

[۱۰] Carroll

[۱۱] Profile likelihood method

[۱۲] Fan

[۱۳] Mammen

[۱۴] Tjøstheim and Auestad

[۱۵] Linton and Nielsen

[۱۶] Back fittinf algorithm of Buja

[۱۷] Hastie and Tibshirani

[۱۸] Amato

[۱۹] Differencing method

[۲۰] Rice

[۲۱] Hall

تفاضلی می توان به برآوردگر کاراتری در مدل رگرسیون ناپارامتری دست یافت. هوروویتس و اسپکواینی[۲۲] (۲۰۰۱) روش تفاضلی را برای آزمون خطی بودن تابع ناپارامتری در مدل رگرسیون ناپارامتری مورد استفاده قرار دادند. اما یاتچیو[۲۳] (۱۹۹۷) اولین بار روش تفاضلی را در مدل های خطی جزئی بکار برد و بر برآورد پارامتر خطی با حذف تابع ناپارامتری بوسیله روش تفاضلی تمرکز کرد. او، همچنین خواص مجانبی برآوردگر پارامتری را با فرض کراندار بودن مشتق مرتبه اول تابع ناپارامتری مورد بررسی قرار داد.

این مجموعه مبتنی بر ۴ فصل است که مطالب هر فصل به اختصار بصورت زیر می باشد:

- فصل اول شامل مقدمه ای بر رگرسیون چندگانه، معرفی برآوردگرها و خواص آن ها در این مدل و روش های رایج برآورد در مدل رگرسیون ناپارامتری است.

- در فصل دوم، پس از معرفی روش تفاضلی، به برآورد واریانس خطا، انجام آزمون در مورد تابع ناپارامتری در مدل رگرسیون ناپارامتری و مدل خطی جزئی، برآورد پارامتر خطی در مدل های خطی جزئی و استنباط در مورد آن با استفاده از روش تفاضلی می پردازیم.

- در فصل سوم، تحت یك محدودیت روی فضای پارامتر خطی و با در نظر گرفتن وجود همبستگی بین ستون های ماتریس طرح، دو برآوردگر تفاضلی برای پارامتر خطی، با فرض همبسته بودن خطاها در مدل خطی جزئی، پیشنهاد شده و تحت تابع زیان متعادل مقایسه خواهند شد.

- در فصل چهارم، تحت شرایط فصل سوم، به مقایسه دو برآوردگر جدید که با روش نیمه پارامتریك در مدل های خطی جزئی بدست آمده اند با استفاده از معیار مخاطره مجانبی، خواهیم پرداخت.

- در فصل پنجم، پس از معرفی مدل های نیمه پارامتریك به ظاهر ناهمبسته، دو برآوردگر تفاضلی با فرض مجهول بودن واریانس جملات خطا پیشنهاد می نماییم.

[۲۲] Horowitz and Spokoiny

[۲۳] Yatchew

قدردانی

سوگند به قلم و آنچه می نویسد. خداوندا با نام تو آغاز می کنم، زیرا از خود توان و نیرویی ندارم و ستایش تو را می گزارم، زیرا ستایش و بندگی تنها در مقابل عظمت و بزرگی تو شایسته است.

در انتها شایسته است که مراتب سپاسگزاری و قدردانی خالصانه و صمیمانه خود را نسبت به استاد عزیزم، جناب آقای دکتر حسینعلی نیرومند که زحمات فراوانی را در جهت هدایت، راهنمایی و همیاری بنده در تهیه و تنظیم این رساله، کشیده‌اند، به جا آورم. از جناب آقای دکتر محمد آرشی، استاد مشاورم، بینهایت ممنون و متشکرم، که با مشاوره‌ها و حمایت‌های بی دریغشان هیچ زمان من را تنها نگذاشته‌اند و با صبر و تلاش ایشان در ایده سازی، هموارسازی و همیاری، توانستم این مجموعه را گرد آوری کنم.

از مدیر دلسوز و پرتلاش گروه آمار دانشگاه فردوسی مشهد، جناب آقای دکتر مجید سرمد به دلیل خدمات، راهنمایی ها و زحماتشان سپاسگزارم.

لازم می دانم کمال تشکر و احترام را نسبت به اساتید داور محترم، آقایان دکتر سیّد محمّد مهدی طباطبایی (دانشگاه فردوسی مشهد)، دکتر احمد پارسیان (دانشگاه تهران)، دکتر هوشنگ طالبی (دانشگاه اصفهان) و خانم دکتر براتپور (دانشگاه فردوسی مشهد)، به دلیل حضور گرمشان، سعه صدرشان در تصحیح این رساله و داوری استادانه شان، به جا آورم.

در طول این کار کوچک از کمکهای مختلف عزیزان دیگری نیز بهره برده ام که لازم می دانم با زبان قاصر خود از آنها نیز قدردانی نمایم. از جناب آقای داوود نژاد و سرکار خانم صادقی که مرا در تهیه برخی مراجع یاری رساندند و همچنین سایر اعضای محترم کتابخانه دکتر فاطمی (دانشکده علوم ریاضی دانشگاه فردوسی مشهد) که اوقاتی مزاحم وقت ارزشمندشان در راهنمایی گرفتن برای یافتن مراجع و کتب مورد نیاز، شدم سپاسگزارم و محبت های بی دریغ مسئولین آموزش بخصوص سرکار خانم محمدی را نیز سپاس دارم.

از منشی گروه آمار، سرکار خانم سلیمانی، مسئولین امور پژوهشی دانشکده علوم ریاضی آقای پازوکی و خانم دامغانی، مدیر امور عمومی آقای مومن، کادر نگهبانی و خدمات دانشکده علوم

ریاضی به دلیل محبت های بی دریغ شان، قدردانی می کنم.

قطعاً این مجموعه بدون نقص نیست، لذا از خوانندگان عزیز تقاضا دارم با دقت موشکافانه خود عیب‌ها و کاستی‌های این رساله را از طریق پست الکترونیکی[1] به اطلاع اینجانب برسانند.

مهدی روزبه – تابستان ۱۳۹۰

---

[1] m.roozbeh.stat@gmail.com

نمادها

$w_j(t_i)$: وزن مشاهده jام در برآورد $f(t_i)$

$\hat{f}^{-i}(t_i)$: معادله رگرسیون ناپارامتری بدون در نظر گرفتن (t_i, Y_i)

$\xrightarrow{D}$: همگرایی در توزیع

$\xrightarrow{p}$: همگرایی در احتمال

$a_n = o(n^{-r})$: $n^r a_n \to ۰$

$a_n = o_p(n^{-r})$: $n^r a_n \xrightarrow{p} ۰$

$a_n = O(n^{-r})$: برای مقادیر بزرگ n دنباله $|n^r a_n|$ همگرا است

$a_n = O_p(n^{-r})$: برای مقادیر بزرگ n دنباله $|n^r a_n|$ در احتمال همگرا است

$N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$: توزیع نرمال n متغیره با میانگین $\boldsymbol{\mu}$ و واریانس $\boldsymbol{\Sigma}$

$\widetilde{\boldsymbol{Y}}$: $\boldsymbol{DY}$

$\bar{\boldsymbol{Y}}$: متغیر پاسخ هموار شده

$\widetilde{\boldsymbol{X}}$: $\boldsymbol{DX}$

$\bar{\boldsymbol{X}}$: ماتریس طرح هموار شده

$\widetilde{\boldsymbol{\epsilon}}$: $\boldsymbol{D\epsilon}$

$\boldsymbol{V_D}$: $\boldsymbol{DVD'}$

$s^۲_{diff}$: برآوردگر تفاضلی مرتبه اول $\sigma^۲_\epsilon$

$s^۲_{LS}$: برآوردگر کمترین مربعات $\sigma^۲_\epsilon$

$s^۲_p$: برآوردگر آمیخته $\sigma^۲_\epsilon$ بر اساس مجموع دو نمونه

$s^۲_{\boldsymbol{D}}$: برآوردگر تفاضلی مرتبه mام $\sigma^۲_\epsilon$

$s^۲_{\boldsymbol{R}}$: برآوردگر محدود شده $\sigma^۲_\epsilon$

$s^۲_{\boldsymbol{RD}}$: برآوردگر تفاضلی محدود شده $\sigma^۲_\epsilon$

$\odot$: عملگر ضرب مستقیم ماتریسی

$\otimes$: عملگر ضرب کرونکر

η: گشتاور غیر مرکزی مرتبه چهارم

$\hat{\boldsymbol{\beta}}_{\boldsymbol{R}}$: برآوردگر محدود شده کمترین توان های دوم $\boldsymbol{\beta}$

$\hat{\boldsymbol{\beta}}_{\boldsymbol{D}}$: برآوردگر تفاضلی کمترین توان های دوم $\boldsymbol{\beta}$

$\hat{\boldsymbol{\beta}}_{\boldsymbol{RD}}$: برآوردگر تفاضلی محدود شده کمترین توان های دوم $\boldsymbol{\beta}$

$\hat{\boldsymbol{\beta}}_{B}$: برآوردگر بیز $\boldsymbol{\beta}$

$\hat{\boldsymbol{\beta}}_{ME}$: برآوردگر آمیخته نااریب کمترین توان های دوم $\boldsymbol{\beta}$

$\hat{\boldsymbol{\beta}}(k)$: برآوردگر ریج کمترین توان های دوم $\boldsymbol{\beta}$

$\hat{\boldsymbol{\beta}}_{G\boldsymbol{D}}$: برآوردگر تفاضلی کمترین توان های دوم تعمیم یافته $\boldsymbol{\beta}$

$\hat{\boldsymbol{\beta}}_{G\boldsymbol{D}}(k)$: برآوردگر تفاضلی ریج کمترین توان های دوم تعمیم یافته $\boldsymbol{\beta}$

$\hat{\boldsymbol{\beta}}_{G\boldsymbol{RD}}$: برآوردگر تفاضلی محدود شده کمترین توان های دوم تعمیم یافته $\boldsymbol{\beta}$

$\hat{\boldsymbol{\beta}}_{G\boldsymbol{RD}}(k)$: برآوردگر تفاضلی محدود شده ریج کمترین توان های دوم تعمیم یافته $\boldsymbol{\beta}$

$\hat{\boldsymbol{\beta}}_{S}$: برآوردگر نیمه پارامتری کمترین توان های دوم $\boldsymbol{\beta}$

$\hat{\boldsymbol{\beta}}_{GS}$: برآوردگر نیمه پارامتری کمترین توان های دوم تعمیم یافته $\boldsymbol{\beta}$

$\hat{\boldsymbol{\beta}}_{GS}(k)$: برآوردگر نیمه پارامتری ریج کمترین توان های دوم تعمیم یافته $\boldsymbol{\beta}$

$\hat{\boldsymbol{\beta}}_{G\boldsymbol{R}S}$: برآوردگر نیمه پارامتری محدود شده کمترین توان های دوم تعمیم یافته $\boldsymbol{\beta}$

$\hat{\boldsymbol{\beta}}_{G\boldsymbol{R}S}(k)$: برآوردگر نیمه پارامتری ریج محدود شده کمترین توان های دوم تعمیم یافته $\boldsymbol{\beta}$

$\hat{\boldsymbol{\beta}}^{SUS}$: برآوردگر $\boldsymbol{\beta}$ در مدل های نیمه پارامتری به ظاهر ناهمبسته

$\hat{\boldsymbol{\beta}}_{F}$: برآوردگر شدنی $\boldsymbol{\beta}$

$\boldsymbol{C}$: $\boldsymbol{X}'\boldsymbol{V}^{-۱}\boldsymbol{X}$

$\boldsymbol{C}_{\boldsymbol{D}}$: $\widetilde{\boldsymbol{X}}'\boldsymbol{V}_{\boldsymbol{D}}^{-۱}\widetilde{\boldsymbol{X}}$

$\boldsymbol{C}_{\boldsymbol{D}}(k)$: $\widetilde{\boldsymbol{X}}'\boldsymbol{V}_{\boldsymbol{D}}^{-۱}\widetilde{\boldsymbol{X}}+k\boldsymbol{I}$

$\boldsymbol{C}_{S}$: $\bar{\boldsymbol{X}}'\boldsymbol{V}^{-۱}\bar{\boldsymbol{X}}$

$\boldsymbol{C}_{S}(k)$: $\bar{\boldsymbol{X}}'\boldsymbol{V}^{-۱}\bar{\boldsymbol{X}}+k\boldsymbol{I}$

$\boldsymbol{M}_{\boldsymbol{D}}(k)$: $\boldsymbol{C}_{\boldsymbol{D}}^{-۱}(k)-\boldsymbol{C}_{\boldsymbol{D}}^{-۱}(k)\boldsymbol{R}'\big(\boldsymbol{R}\boldsymbol{C}_{\boldsymbol{D}}^{-۱}(k)\boldsymbol{R}'\big)^{-۱}\boldsymbol{R}\boldsymbol{C}_{\boldsymbol{D}}^{-۱}(k)$

$\boldsymbol{M}_{S}(k)$: $\boldsymbol{C}_{S}^{-۱}(k)-\boldsymbol{C}_{S}^{-۱}(k)\boldsymbol{R}'\big(\boldsymbol{R}\boldsymbol{C}_{S}^{-۱}(k)\boldsymbol{R}'\big)^{-۱}\boldsymbol{R}\boldsymbol{C}_{S}^{-۱}(k)$

$\kappa(\boldsymbol{A})$: عدد شرطی $\boldsymbol{A}$

فهرست مطالب

لیست جداول

فصل ۱

مدل های رگرسیون چندگانه و ناپارامتری

۱.۱ مقدمه و تاریخچه

رگرسیون چیست؟

در کتب آماری رگرسیون به عنوان روش تعیین و تحلیل روابط نادقیق بین متغیرهای آماری تعریف می شود. بنابراین در رگرسیون دو مطلب مورد توجه قرار می گیرد:

- تعیین روابط بین متغیرها
- تحلیل روابط بدست آمده

کلمه رگرسیون در لغت به معنای ''برگشت'' است و ظاهرا تصویر روشنی از آنچه که رگرسیون آماری انجام می دهد، ارائه نمی کند. دلیل این نام گذاری به مطالعات گالتون در مورد رابطه بین قد پسر و قد پدر مربوط می شود. گالتن با رسم نمودار قد پسر در مقابل قد پدر (نمودار پراکنش) بر اساس حجم بزرگی از داده ها به این نتیجه رسید که اغلب افراد بلند قد، پسرانی کوتاه قدتر و اکثرا افراد کوتاه قد، پسرانی بلند قدتر از خود دارند و این پدیده را ''برگشت به میانگین[۱] '' نامید و از آن پس نام ''رگرسیون'' بر روش های بررسی روابط بین متغیرهای آماری باقی ماند. تعیین روابط بین متغیرهای آماری را رگرسیون توصیفی و تحلیل این روابط را رگرسیون استنباطی می نامند. بعنوان مثالی از رگرسیون، فرض کنید می خواهیم بدانیم که آیا مصرف سرانه سیگار با

[۱] Regression to the mean