بسم الله الرحمن الرحیم

قالوا سبحانک لا علم لنا الا ما علمتنا انک انت العلیم الحکیم

گفتند منزهی تو ما را دانشی نیست جز آنچه تو
بما آموختی همانا توئی دانشمند حکیم

(قرآن آیه ۳۲ سوره بقره)

رگرسیون های نـرم L_1 گسسته و پیوستـه

پیشنهاد الگوریتم های تقریب گسسته وبرازش پیوسته رویه تمرکز

توسط

بیژن بیدآبـاد

پایان نامـــه

بعنوان بخشی از پیش نیازهای دریافت درجه

دکترای تخصصی (Ph.D)

در رشتـه

اقتصـاد

دانشگاه آزاد اسلامـــی

تهران،ایران

باهمکاری دانشگاه نوشاتل سوئیـس

بادرجه ......عالی... ارزیابی وتصویب شد.

دکتر محمدجعفرمجرد	استاداقتصاد (راهنما)	....
دکتر اسدمنجمـــی	استادآمـار (مشاور)	....
دکتریداللـــه دوج	استاد آمار وتحقیق در عملیات دانشگاه نوشاتل سوئیس (مشاور)	....
دکتراکبرکمیجانـــی	استاد اقتصاد (عضوکمیته)	....
دکترمحمدرضاشریف آزاده	استاد اقتصاد (رئیس گروه اقتصاد)	....

دی ماه ۱۳۶۸

" سایرکاربردهاوویژگیهای این روش [میانه های جمعی] بی تردیدزمانـــی روشن خواهدشدکه متخصصین توجه خودرابه این مسئله معطوف سازند۰۰۰۰۰۰ پذیرش بی چون وچرای روش حداقل مربعاتموجه نیست۰۰۰۰۰۰۰ راســـــت اندیشی نبایدقبل ازوحی بدعتی بگذارد۰

ف ۰ ی۰ اجورث ۱۸۸۸

(Farebrother(1987))

چکیـــــده

دراین رساله یك مرورنسبتا " کامل ازتحلیل های آماری برمبنای نرم L_1 وروشهـــای مربوطه ارائه می شودکه امیدمی رودجای خالی چنین بررسی رادرزمینه برآوردنـــرم L_1 پرکند. الگوریتم های نزولی متفاوتی برای برآوردنرم L_1 پارامترهای رگرسیون هـــای ساده وچندمتغیره پیشنهادمی شوندکه درمیان الگوریتم های موجودازصحت وکارآئـــی خوبی برخوردارهستند. باگریزی به رویه تمرکزوبسط برازش نرم L_1 پیوسته ، بـــرآورد فرم های تابعی منحنی لورنزموردملاحظه قرارمی گیرد. ازاین بابت اطلاعات توابع چگالـی احتمال پیوسته که ازنوع داده های پیوسته می باشنداستفاده می شوند.

" این رساله اززبان انگلیسی به فارسی بــرگـــر دانــده شــده است "

سپاسگزاری

این سپاسگزاری هرگز دیون مرا به افرادی که در این مطالعه مرا همراهی کردند، جبران نخواهد کرد. به هر حال، باید یادآور یاری آنان شوم که بدون آنها هرگز موفق به تکمیل پیش نیازهای دریافت درجه دکترای تخصصی در اقتصاد نمی شدم.

در ابتدا از دکتر عبدالله جاسبی رئیس دانشگاه آزاد اسلامی و دکتر محمدرضا شریف آزاده رئیس گروه اقتصاد تشکر می نمایم که یك بورس چهارساله کامل برای تحصیل در این دانشگاه و همچنین کمك های فراوانی جهت شرکت در کنفرانس های ملی و بین المللی به اینجانب اعطا نمودند. تشکرات صمیمانه خود را خدمت استاد راهنمایم دکتر محمدجعفر مجرد که با راهنمائی و تشویق خود در تمام مراحل این تحقیق و با صرف اوقات زیاد سبب به ثمر رسیدن این کار گردیدند و همچنین به اساتید مشاور خود دکتر احمد منجمی و دکتر یدالله دوج از دانشگاه نوشاتل سوئیس برای توصیه ها و پیشنهادات ارزشمند ایشان و به دکتر اکبر کمیجانی عضو کمیته رسیدگی برای بحث ها و توصیه های سازنده ایشان تقدیم می نمایم. همچنین از اساتید گروه اقتصاد که اوقات زیادی در رابطه با مسائل آموزشی اینجانب صرف نمودند سپاسگزاری می کنم.

از دکتر یدالله دوج از دانشگاه نوشاتل سوئیس و برگزارکننده "**اولین کنفرانس بین المللی تحلیل های آماری بر مبنای نرم L_1 و روشهای مربوطه**" بخاطر میهمان نوازی ایشان در طول مدت شرکت من در این کنفرانس و تخصیص وقت برای ارائه مقالاتم و دکتر کریم صدیقی از دانشگاه شیراز بخاطر دعوت برای ارائه دو سخنرانی در "**سومین سمینار آنالیز ریاضی**" و همچنین دکتر M.McAleer از دانشگاه ملی استرالیا برای وقتی که به من برای ارائه مقاله ام در "**گردهمائی استرالیا - آسیا جامعه اقتصادسنجی**" دادند و از دکتر H.Imai و M. Iri از دانشگاه چوب بخاطر دعوت به "**سیزدهمین سمپوزیوم بین المللی برنامه ریزی ریاضی**" سپاسگزاری می نمایم.

همچنین از استاد برجسته و مشهور علم آمار C.R.Rao بخاطر بحث، راهنمائی و تشویق ایشان و از آقایان دکتر ها فریدون اهرابی، R.W.Farebrother, P.Bloomfield, H.Spath, S. Portnoy , M.R.Osborne , S.C.Narula , R. Koenker که دو تا از مقالات اینجانب که در کنفرانس نوشاتل ارائه گردید را مورد بحث قرار دادند تشکر

می نمایم.

ازاساتیدزیرازدانشگاه‌های مختلف دنیاکه کتب ومقالات خودکه مربوط به‌این تحقیق بودرابرایم فرستادندسپاسگزاری می نمایم.

J.Antoch,P.Bloomfield,R.H.Bartels,H.Brunk,G.Bassett,D.Bradu, M. Csorgo,P.H.Calamai ,S.S.Chiu,A.R. Conn,T. M. Cavalier, Y.Dodge, J. Dupacova,R. Dutter, J.r. Eriksson,A.K.md.Ehsanes Saleh, H . Ekblom, R.M.Freund, R. Fourer,R.Fletcher, V.V.Fedorov,R.W.Farebrother, R.Gonin,J. Galpin, D.Goldfarb,S.A.Gustafson,E.Gorecki,J. L.Goffin,C.C.Gonzaga,W.Heiser,J.Hald,M.L.Hart,W.Hardle,T.Hettmansperger,M.Iri,H. Imai,J.P.Ignizio,K.Jittorntrum,K.Jafarpour,J . Jureckova,L.Kaufman,K.O.Kortanek,N.Kaergard,M.Kojima,J.S.Marron, A.Money,A. Marazzi,N.Megiddo,C.McConnell,S Mizuno,J.Mitchell, J. Mckean , J.A.Menendez, S.C. Narula,M. R .Osborne, U.Peters,P. Pilibossian , C.Roos,P.Rousseew,F.Reza,J.Renegar,R.Rossi,S.M.Stigler, D.L.Souvaine,G.Stangenhaus,V.Sposito,Su.Chun,R.I.Shrager, A.F. Siegel,H.D.Sherali,H . Spath, A.H.Seheult,E.Seneta,S.J.Sheather, A. Schrijver,L . Seiford,W.Steiger,I. M.Steele,R.M.Schrader,D.S.Tracy, T. Taguchi,E. Trauwaert, J.A.Tomlin,K.Turner,M.j.Todd,S.H.C.du Toit, J.p.Vial,R.J. Vanderbai, A.H.Welsh, G.Wesolosky, C.S.Withers G.A.Watson,R.S.Womersley,U.Zimmerman.

تشکرات صمیمانه وقلبی خودم راخدمت پدرومادرعزیزومهربانم آقای محمدعلی بیدآبادوخانم ملیحه رابطی بواسطه فداکاری های همیشگی ایشان که‌با مشکلات حادوعدیده زندگی هرگزازتعلیم وتربیت فرزندان غافل نشدندتقدیم می دارم ازخداوندمتعال خواهانم که سایه ایشان رابرسرخانواده مستدام دارد. ازهمسرعزیزم نادره‌راستین بخاطرتامین محیطی مناسب برای مطالعه وتحقیق وبعهده گرفتن امورزندگی درطی این مدت صمیمانه سپاسگزارم. ازبرادربزرگم دکتربهروزبیدآبادازبابت مشورت هاوبحث های ارزنده‌ایشان تشکرمی نمایم . برادروخواهرکوچکم پیروزوفرانک بیدآبادودخترم شادی بانوبخاطرتحصیل من که‌فرصتی برای رسیدگی به‌اموراایشان راباقی نمی گذاشت متشکرم .

ازآقایان مسعودروغنی زنجانی رئیس ،محمدجعفراسلامی معاون امورمناطق ومحمد حسن فولادی مدیردفتربرنامه‌ریزی منطقه‌ای سازمان برنامه‌وبودجه بخاطردراختیارگذاشتن تسهیلات کامپیوتری تشکرمی نمایم٠ازآقایان دکترمحمدنیری وندمعاون آموزشی و دانشجوئی ،محمدحسین سدهی نیامعاون مالی واداری ،محمدفامیلی مسئول دایره اداری ، حسین انصاری آملی مسئول حسابداری آقای هاشم اخلاقیان کارمنددایره اداری ،خانم فاطمه بنازاده مسئول مرکزکامپیوتر ،خانم لیندا‌خدابخش مقدم مسئول کتابخانه ،آقای علی رضائی خانم فرحدخت شریل لی واعظم واثقی کارکنان کتابخانه،آقایان بهروزگندمکار ،پرویز طاهری ،کیومرث ملکی مسئولین امورتدارکات ،وخانمهانوشین سرشار ،نیماابریشم کار ورزیتاعموئی کارکنان دبیرخانه وآقایان قدرت الله امانی وغلامحسین قایقی وسایر کارکنان دوره عالی تحقیقات (دکتری) دانشگاه‌آزاداسلامی که هرکدام بنحوی دررفع مشکلات مربوط به‌اینجانب دلسوزانه تلاش نمودندسپاسگزاری می نمایم ٠ اززحمات خانم رویانقوی جورشری وآقای فیروزفراشی اوغانی که تایپ نسخه فارسی رابعهده‌داشتند تشکرمی نمایم٠

بیژن بیدآباد

۱۳۶۸/۱۰/۳

فهرست مطالب	شماره صفحه

فهـــرســـتمطـــالـــب | شماره صفحه

فهرست مطالب | شماره صفحه

فهرست مطالب	شماره صفحه
فصل پنجم برازش نرم L_1 پیوسته ومنحنی لورنز	۱۲۳
۱- مقدمه	۱۲۵
۲- نرم L_1 توابع پیوسته	۱۲۵
۳- برازش نرم L_1 یك پارامتری خطی	۱۲۶
۴- برازش نرم L_1 دوپارامتری خطی	۱۲۸
۵- رویه تمرکز	۱۳۰
۶- تقریب نرم L_1 پیوسته منحنی لورنز	۱۳۵
فصل ششم خلاصه ونتایج ،توصیه برای تحقیقات بعدی	۱۴۳
۱- خلاصه ونتایج	۱۴۵
۲- توصیه برای تحقیقات بعدی	۱۴۷
ضمائم	۱۴۹
ضمیمه الف - برنامه‌کامپیوتری برای مسئله میانه‌وزنی	۱۵۱
FUNCTION LWMED	۱۵۱
ضمیمه ب - برنامه‌کامپیوتربرای الگوریتم ۲	۱۵۲
PROGRAM BL1S	۱۵۲
ضمیمه ج - برنامه کامپیوتربرای الگوریتم ۴	۱۵۳
PROGRAM BL1	۱۵۳
SUBROUTINE COL1	۱۵۴
SUBROUTINE COL2	۱۵۴
SUBROUTINE COL3	۱۵۴
کتابشناسی	۱۵۵

فصـــــــل اول

" مقـــدمــــه عمـومـی "

۱۰ مقـــدمـــه

گرچه‌روش برآوردحداقل‌مربعات پارامترهای‌رگرسیون معمولترین روش‌مورداستفاده‌است،در سالهای‌اخیرروشهای‌جایگزین دیگری‌موردتوجه‌بسیارقرارگرفته‌اند.طبیعتا" توجه‌به‌سایرروشهای برآورددراثرعملکردنامطلوب برآوردکننده‌حداقل‌مربعات‌درحالات‌خاصی‌که‌بعضی‌ازفرضیات الگو برقرارنیستندویازمانی که درمیان رگرسورهاهمبستگی زیادی وجودداردایجادشده‌است . به هرحال رگرسیون حداقل مربعات درخیلی حالات غیرگوسی (non-Gaussian) ازبهینه دوراست ،مخصوصا" زمانیکه خطاها ازتوزیع‌هائی بادم طولانی پیروی می کنند. مشخصا" ، هرگاه واریانس‌خطاهابی نهایت است . گرچه شم ریاضی انسان ممکن است ملاحظه خطاها باواریانس‌بینهایت راناپذیردولی درحالات متعدد،مطالعات نشان‌داده‌اندکه درواقع توزیع‌های خاصی باواریانس‌بینهایت می توانندالگوهای بسیارمناسبی باشند. یك واریانس بینهایت به معنی توزیع‌خطای دم‌کلفت باتعدادی نقاط پرت می باشد. البته ، توزیع ملاحظه شده متغیرهای اقتصادی هیچگاه واریانس‌های بینهایت رانشان‌نمی دهند.به‌هر حال ،موضوع مهم این‌نیست که گشتاوردوم عملا" بینهایت باشد،امادامنه بین دهك‌ها در رابطه بادامنه بین چارك‌هابه‌اندازه کافی بزرگ است که‌میتواندتوجیه کننده‌واریانس بینهایت باشد. اگراکثریت خطاهادرالگوازتوزیع نرمال‌تبعیت نمایند،اغلب اتفاق می افتد که‌تعدادکمی ازمشاهدات ازتوزیع‌متفاوتی باشند.بدین معنی که نمونه آلوده به نقاط پرت باشد. ازآنجائیکه حداقل مربعات وزن زیادی به نقاط پرت می دهدشدیدا" وابسته به‌نمونه می شودوکاملا" مشخص است که عملکرداین برآوردکننده دراین حالت بطورقابل ملاحظه‌ای نزول می نماید. حتی اگرخطاهاازتوزیع‌نرمال پیروی کنننددراین حال هم جایگزین حداقل مربعات لازم بنظرمی رسد. مخصوصا" اگرفرم الگودقیقا" شناخته‌نشده باشدویاهرنوع دیگری ازخطای تشخیص وجودداشته باشد. پیش ازاین ، اگرتابع‌زیان‌کوادراتیك معیار مناسبی اززیان نباشد،حداقل مربعات خیلی رضایتبخش نخواهدبود. زیان مشخص کننده‌جدی بودن خطای پیش بینی غیرصفربرای محقق است که خطای پیش بینی اختلاف بین‌مقادیر مشاهده شده وپیش بینی شده متغیرپاسخ می‌باشد. برای برخی مسائل اقتصادی نشان داده شده است که حداقل قدرمطلق خطاهانتایج بهتری تاحداقل مربعات می دهدزیراولی کمترازدومی به خطاهای بزرگ حساس تراست ونتیجتا" به‌نقاط پرت مقاوم است . باید

خاطرنشان ساخت که اگرخطاها ازتوزیع لاپلاس پیروی کنندبرآوردهای حداقل قدرمطلق خطا خواص حداکثراحتمال رادارندوبنابرآن بطورمجانبی کاراهستند.

(1757) Boscovich اولین فردی بودکه برآوردکننده حداقل قدرمطلق خطا راپیشنهادمی نماید.(1779) Laplace مسئله رابرای الگوی یك پارامتری ساده‌حل کرد.(1888) Edgeworth یك الگوریتم برای این مسئله حداقل سازی معرفی نمود (1958) Karst و (1971) Sharpe ، (1972) Rao,Srinivasan و (1973) Appa,Smith و (1976) Smith و Sposito و (1983) Sposito و Josvanger الگوریتم های تکراری برای برآوردپارامترهای رگرسیون خطی ساده را توسعه دادند.

درحالت رگرسیون خطی چندمتغیره ،(1930) Rhodes و (1940) Singleton پیشنهاد Edgeworth رابسط می دهند. به‌هرحال این روشها با افزایش تعدا متغیرها مهارناپذیرمی شوند.(1955) Ferguson و Cooper و Charnes نشان دادندکه رگرسیون حداقل قدرمطلق خطاها اساسا " یك مسئله برنامه ریزی خطی است و (1959) Wagner آنرا بعنوان یك مسئله برنامه ریزی خطی فرموله می کند. روش سیمپلکس تعدیل‌شده (1966) Young و Barrodale وروش نزولی (1967) Usow وروش برنامه ریزی فاصله‌ای (1969) Ben- Israel و Robers متعاقبا " مطرح شدند(1971) Abdelmalek این برآوردرا با حل دنباله‌ای ازمسائل غیرخطی بدست می آوردو (1973) Schlossmacher آنرا باروش حداقل مربعات وزنی تکراری‌محاسبه کرد.(1973) Roberts و Barrdale یك الگویتم خیلی کاراویژه برای این مسئله معرفی می نماید. الگوریتم آنان چندین دورتکرارسیمپلکس رادریکی ادغام می نماید.(1979) Kung و Frome و Armstrong تجزیه LU (ماتریس های مثلثی پائین وبالا) رابرای معرفی یك الگوریتم سیمپلکس تجدید نظرشده برای این مسئله استفاده می نماید.(1980) Steiger و Bloomfield (1981) Wesolowsky نیزدوالگوریتم نزولی طرح می نمایند.

مطالعات مقایسه ای معدودی وجوداردکه عملکردالگوریتم های مختلف راباویژگی‌های مختلف مسئله همانندهمخطی کامل ، مشخص کردن توزیع‌های مختلف خطاوکارآئی

محاسباتی وفضای حافظه لازم وزمان اجرای کامپیوتررا می آزمایند. به هرحال ،نتایج نشان میدهندکه هرالگوریتم درشرایط خاصی عملکردخوبی داردویك الگوریتم منحصربفرد وجودنداردکه تمام خواص خوب رادارا باشد. میتوان نتیجه گرفت که یك شکاف محاسباتی بین برآوردکننده های حداقل مربعات وحداقل قدرمطلق خطاها وجوداردوتلاش برای حذف این شکاف ارزشمندمی باشد.

اگرچه برآوردکننده حداقل قدرمطلق خطا خیلی قدیمی است ،اخیرا "دردودهه اخیربدلیل خواص نامطلوب حداقل مربعات بمتون علمی واردشده وجلب توجه مینماید. حالا این روش در کتب درسی اقتصادسنجی همانند(1986) Kmenta و(1977) Maddala موردبحث قرار می گیرد. رساله های فوق لیسانس ودکتری متعددی دراین موضوع ودردپارتمان های مختلف مانند(1961) Lawson و(1965) Burgoyne (1965) Gentleman و(1967) Barrodale و(1968) Ovsen و(1969) Lewis و(1970) Cline و(1970) Hunt و(1971) Groucher و(1972) Henriksson و(1973) Bassett و(1974) Forth و (1975) Anderson و(1977) Ronner و(1981) Nyquist و (1981) Clarke و(1981) Kotiuga و(1983) Gonin و(1985) Busovaca و( ) Kim که اخیرتر ازبقیه می باشند(نگاه کنیدبه کتابشناسی برای دپارتمان ها ودانشگاه های مربوطه) دراین موردنوشته شده است .

خواص نیرومنداین برآوردکننده مزیت آن برای کارکردن باتوزیع های خطابا واریانس زیادمی باشد ،ازآنجائیکه پدیده های اقتصادی زیادی مانندتوزیع درآمدشخصی ،سودسهام ، قیمت های سفته بازی ،قیمت موجودی وکالاها ،اشتغال ،اندازه دارائی بنگاه های تجاری ، معادلات تقاضا ،نرخ بهره ،جریان وجوه خزانه ،بیمه ،انتظارات قیمت وخیلی دیگر از متغیرهای اقتصادی درون گروه واریانس بینهایت می افتند(نگاه کنیدبه (1972) Granger , Orr و(1977) Nyquist,Westlund و(1965) Fama و(1981) Goldfeld , Quandt و(1971) Sharpe) جلب توجه اقتصاددانان به این برآوردکننده لازم بنظرمی رسد. کارهای متعدددیگری همانند تحلیل تقاضای بین صنایع (1959) Arrow , Hoffenberg ، الگوی

سرمایه‌گذاری(1987) Meyer , Glauber ، تحلیل های سهـــــــام و اوراق بهادار(1971) Sharpe ،تحلیل های سرمایه گذاری دانمـــــــارك (1987) Kaergard ونظائرآن وجوددارندکه برتری برآوردکننده حداقل قدرمطلق خطاهـــــا به حداقل مربعات راتائیدمی نمایند٠

نظریه های جدیدومتعدداقتصادی فرضیه عقلائی بودن رفتارانسانی راسست می کننــــد٠ این ناعقلائی نسبی منبع مهمی درواریانس‌های زیادونقاط پرت درداده‌های اقتصـــــادی می باشد،بنابراین برآوردکننده حداقل قدرمطلق خطاهادرحالاتی که عقلائی بودن فرضـــی قوی می باشدبرآوردکننده‌مناسبی است ٠

کاربردوسیع‌دیگراین برآوردکننده برروی داده‌های آماری باخطاهای اندازه‌گیری اسـت ٠ این نوع خطاهاباعث واریانس‌زیادشده وسبب می شودکه مشاهدات دورازمحل حقیقی خــود واقع شوندوبوضوح نقاط پرت راایجادنمایند٠وجوددونوع مهم ازخطاهای اندازه‌گیری خطـــای نمونه‌ای وغیرنمونه‌ای ،مشخصا " درکشورهائی که دارای آمارضعیف هستندهمانندکشورهای درحال توسعه‌این برآوردکننده رایك ابزارتحلیل اساسی می نماید٠

خطاهای تشخیص ناشناخته درالگوهای رگرسیون به دلیل پیچیدگی‌رفتارانسان همیشـــه درفرمولاسیون ریاضی مسائل مرتبط باانسان اتفاق می افتد٠ خطای تشخیص درزمانیکــــه فرمولاسیون معادله رگرسیون یایکی ازفروض آن ناصحیح باشداتفاق می افتد٠ دراین مقولــه زمانی که یك فرض تئوری رفتاری مربوطه یافرمولاسیون الگولغومی شودیایك متغیرتوضیحی مناسب حذف ویایك متغیرنامناسب اضافه می شود،تغییرکیفی متغیرتوضیحی نادیـــــده انگاشته می شودیافرم ریاضی رگرسیون ناصحیح انتخاب می شودیاتشخیص ناصحیح گونـه‌ای که خطاواردمعادله رگرسیون می شودبکار برده‌می شودویانظائرآن خطای تشخیص وجـــــود دارد(همچنین نگاه کنیدبه (1986) Kmenta) ٠ ازآنجائیکه خطاهای تشخیـــص همیشه برای محقق روشن نیستند،حداقل مربعات یك برآوردکننده ضعیف می باشدوسایـــــر گزینه‌هامانندحداقل قدرمطلق خطاهاجذاب می شوند٠

اگرچه برآوردکننده حداقل قدرمطلق خطاهاازخواص بهینه‌ای درخیلی ازمسائل اقتصـــادی برخورداراست یك ابزارمعمول نیست ٠ تااندازه‌ای این مسئله مربوط به مشکلات محاسباتـی باتوابع قدرمطلق می باشد٠زمانیکه الگوبزرگ می شودومعادلات بطورهمزمان واردمی شوند،

مشکلات محاسبه افزایش می یابد٠ مسئله دیگری که در این برآوردکننده وجوددارد خصوصیات فضای جواب است که کاملا" شناخته نشده و فرم بسته مربوطه هنوز بدست نیامده است٠

بدین ترتیب این سه مسئله مهم فرم بسته (جبری) مشکلات محاسباتی و خواص فضای جواب بزرگترین موانع برسر راه استفاده عموم از برآوردکننده نرم L_1 می باشند٠ هر تلاشی در رفع این موانع میتواند ارزشمند باشد٠ یکی از اهداف این تحقیق بهبود این مسائل است٠

از آنجائیکه فرم بسته برآوردکننده نرم L_1 هنوز بدست نیامده، سعی می شود که با بکارگیری تکنیک مشتق گسسته در جهت مشتق گرفتن از تابع ابژکتیو نرم L_1 دیدگاه جدیدی را ارائه دهیم٠ این مشتق گیری برروی متغیرهای با دامنه گسسته همراه با مشتق گیری معمولی بر متغیرهای با دامنه پیوسته دانش ما را در مورد فرم بسته جبری مسئله افزایش خواهد داد٠ برای بهبود صحت، سرعت و بطورکلی کارآئی محاسبه برآوردکننده نرم L_1 چهار الگوریتم پیشنهاد می شود که دوتا از آنها برای الگوی رگرسیون ساده و دوتای دیگر برای رگرسیون چندمتغیره می باشند٠ با بررسی خواص الگوریتم های پیشنهادی، خصوصیات زیادی از فضای جواب روشن خواهد شد٠

به هرحال، کار در مورد برآورد نرم L_1 گسسته خاتمه نمی یابد و سعی میکنیم موضوعات فوق را در مورد حالت پیوسته برآورد نرم L_1 تعمیم دهیم، اگرچه، حالت پیوسته برآورد نرم L_1 تابحال در اقتصادسنجی بکار گرفته نشده است، این تکنیک برای حل یکی از مسائل مشهور اقتصادسنجی که عبارت از برآورد منحنی لورنز باشد بکار گرفته خواهد شد٠

توزیع درآمد فردی اغلب برروی منحنی لورنز تصویر می شود٠ منحنی لورنز که یك حالت ساده از رویه تمرکز است از معمول ترین روشها برای توصیف توزیع درآمد می باشد٠ توزیع درآمد یکی از قدیمی ترین پدیده هائی است که اقتصاددانان سعی داشته اند که آنرا از طریق الگوسازی ریاضی بیان نمایند٠ (Cramer (1973 در کتابش این مسئله را تحت عنوان " توزیع درآمد فردی ٠ ٠ ٠ یك مسئله حل نشده " مورد بحث قرار می دهد٠ با روشن نمودن چولگی توزیع درآمد، وی بحث می نماید که این خصوصیت قویا " در جوامع مختلف و زمانهای متفاوت دیده می شود، ولی علیرغم این شرط دانسته، یك الگوی قابل قبول عمومی وجود ندارد و این موضوع هنوز بحث انگیز است٠ بحث می شود که

توزیع‌های لوگ ـ نرمال وپارتودوتابع رقیب برای توضیح این پدیده‌می باشند ،هردوی این توزیع‌هارا میتوان به داده‌های آماری درآمدهرجامعه‌ای برای بدست آوردن پارامترهای توزیع مربوطه برازاندواین پارامترهارا بعنوان معیارهای نابرابری موردا ستفاده قرارداد.به هرحال(1973) Cramer نتیجه‌گیری می نماید .که‌توزیع پارتو برازش خوبی برای صاحبان درآمدبالاتربوده ،ولی لوگ ـ نرمال توزیع خوبی برای دامنه کامل درآمدی می باشد .

توزیع پارتوهمیشه‌واریانس محدودندارد .مقادیرمعین پارامترپارتوسبب بینهایت شدن واریانس توزیع می شود .همانطورکه قبلا" بحث شدهرتوزیع ملاحظه شده درآمد دارای واریانس‌محدودمی باشد .لزومی نداردکه‌نگران نمونه‌باشیم ،تنها سئوال این است که‌آیا نمونه ازیك‌توزیع با واریانس محدودیانا محدودبدست آمده‌است یانه(Mandelbrot(1960,61 این نکته توزیع پارتورا موردنظرقرارداده وسعی می نمایدمفهوم واریانس بینهایت رابا شکل جانبی دم های بالای توزیع‌های درآمدملاحظه شده‌توجیه نماید .بدین معنی که‌فاصله‌بین دهك هابه فاصله بین چارك هابه‌اندازه‌کافی بزرگ است که بتوانیم بگوئیم واریانس‌بینهایت است. به‌هرحال وی سعی بریافتن یك توزیع باثبات (برای خواص این گونه توزیع‌هانگاه‌کنید به (1977) Kendall, Stuart) با میانگین محدود ،واریانس‌بینهایت وحداکثر چولگی به‌راست دارد . دیدگاه وی برای تاکیدبرتوزیع‌های باثبات بدلیل خواص زیبای اینگونه توزیع‌ها می باشد . بدین معنی که ،توزیع احتمال ترکیب خطی n متغیرتصادفی مستقل بایك توزیع‌احتمال مشخص ازهمان نوع‌توزیع متغیرهای تصادفی اولیه باشد .کل درآمدمجموع تمام درآمدهاازمنابع مختلف می باشدوازآنجائیکه تمام مولفه‌های درآمدی دارای همان نوع توزیع‌کل درآمدی می باشند، توزیع‌های باثبات برای این مسئله مناسب بنظرمی رسند . به‌هرحال ، (1960 , 61) Mandelbrot ملاحظه نمودکه توزیع‌پارتو ـ لوی (Pareto - Levy) سه نکته وی رابرقرارمی نماید .متاسفانه ،تابع‌توزیع پارتو ـ لوی وهمچنین تابع‌چگالی آن رانمی توان به یك فرم تحلیلی ساده نوشت وبایدتوسط تبدیل لاپلاس یاتابع مفسربیان شود .

دیدگاههای دیگری نیزوجوددارندکه توزیع‌درآمدرافرموله می نمایندونهایتا" به توزیع پارتوختم می شونددراین مقوله ،میتوان به‌الگوی (1953) Champernowne

اشاره‌کردکه‌ازفرآیندتصادفی مارکف (برای توضیحاتی درمورداین فرآیندنگاه‌کنیدبه Bidabad , Bidabad (1989c,d) استفاده‌می نماید. (1955) Simon و(1965) Steindle یك متغیرتصادفی گسسته را ملاحظه‌می نمایند که دارای توزیع وضعیت - یکنواخت بادم پارتومی باشد. پیش ازاین تاکیدمابرتوزیع درآمد فردی بود، مطالب فوق همچنین درموردسایرتوزیع‌های اندازه که‌میتواندیك متغیرساکن هماننددارائی هاوثروت ویایك متغیرجریان همانندتولیدوفروش برای هرواحداقتصادخرد مانندفردیابنگاه باشد(نگاه کنیدبه (1973) Cramer).

تکنیك رویه تمرکزروشی است برای تحلیل توزیع‌های چوله وازاین بابت به‌توزیع‌های پارتوولوگ - نرمال مرتبط می شود. یك رویه تمرکزساده‌منحنی لورنزنامیده‌می شودکه مشخصا " برای توضیح نابرابری توزیع‌درآمدابداع شده‌است . همانطورکه توسط (1977) Stuart , Kendall موردبحث قرارمی گیرد، رابطه‌ای بین‌مساحت زیرمنحنی لورنز وتابع‌توزیع احتمال جامعه آماری مربوطه وجوددارد. بدین معنی که، هرگاه تابع‌توزیع‌احتمال دانسته باشد، می توانیم ضریب جینی (Gini) مربوطه‌رابعنوان معیارنابرابری پیدا کنیم . ولی همانطورکه (1975) Hagen و(1989 a,b) Bidabad,Bidabad ابرازمی نمائید، دومنحنی لورنزمتفاوت باانحناهای متفاوت میتوانندنسبت های جینی یکسانی رادهند. بنابراین ، ممکن است بعضی توزیع‌های‌محدود درآمد اتفاق افتدکه‌نسبت جینی راتغییرندهد. این نکته قابلیت مقایسه دوضریب جینی رابعنوان اندازه نابرابری ضعیف می سازد. بدین ترتیب بایدبه‌عقب برگردیم وفرم های تابعی منحنی های لورنز را برای تعبیرنابرابری توزیع‌درآمدوهمچنین تغییرات آن استفاده‌نمائیم.

برآوردفرم‌تابعی منحنی لورنزنیزبا مشکلات دیگری مواجه‌است . برای این برآورد، باید فرم تابعی مناسبی راکه قابلیت قبول انحناهای مختلف راداشته باشدتعریف نمائیم. این نکته توسط (1989 a, b) Bidabad و Bidabad مورد بحث قرارمی گیرد. دراین مقالات ، طبیعت اتورگرسیوخطاهادرداده‌های انباشته لازم برای برآوردمنحنی لورنزنیزروشن می شود. مسئله دیگری وجودداردکه برای ایجادمجموعه داده‌های آماری لازم برای برآوردپارامترهای منحنی لورنزمربوطه ، حجم زیادی ازمحاسبات بر روی داده‌های درآمدخام نمونه اجتناب ناپذیراست . بوضوح ، این مسائل علیرغم مشکلات

محاسباتی آنها سبب ضعیف نمودن برآوردپارامترها می شود. برای اجتناب ازاین موضوع سعی براین داریم که فرم تابعی منحنی لورنز را با استفاده از داده‌های پیوسته برآورد کنیم. دراین تحقیق تابع چگالی احتمال درآمد جامعه آماری را برای حصول مقادیر برآوردپارامترهای تابع لورنز بکار می بندیم. روش برازش نرم L_1 پیوسته که برای برآوردپارامترهای رگرسیون بسط داده خواهد شد برای حل این مسئله بکار گرفته می شود. به هرحال، بیشتر بر روی دو تابع چگالی احتمال رقیب پارتو و لوگ - نرمال تمرکز خواهیم کرد. از آنجائیکه اولی بسادگی انتگرال پذیر است، مسئله کلی برای حصول تابع لورنز مربوطه وجود ندارد و تابع به صورت یکتا بدست می آید. ولی در مورد دوم، تابع چگالی لوگ - نرمال که عملکرد بهتری برای دامنه درآمدی کامل دارد تا توزیع پارتو، انتگرال پذیر نیست و نمی توانیم تابع لورنز مربوطه اش را تعیین نمائیم. از این بابت باید مسئله را با تعریف فرم تابعی منحنی لورنز کلی و با بکار گیری برازش نرم L_1 برای برآوردپارامترهای مربوطه حل نمائیم.

قسمت های اصلی این تحقیق به شکل زیر است این فصل بعنوان نقطه شروع با مقدمه‌ای بر فضای نرم وخواص آن ادامه می یابد. هدف این قسمت آشنا کردن خواننده با علائم و ویژگی های نرم می باشد. رابطه نرم و رگرسیون موضوع قسمت بعد است. دراین قسمت با اشاره بر رگرسیون نرم L_1 بعنوان حالت خاصی از رگرسیون نرم L_p، موقعیت آن در رگرسیون‌های نرم از یک دیدگاه آماری ملاحظه می شود. در قسمت بعد خواص مهم برآوردنرم L_1 بحث می شوند. خواص ریاضی رگرسیون حداقل قدر مطلق خطاها مانند خاصیت تغییرناپذیری، تحدب شرط بهینه‌گی، حالات یکتا و غیریکتا بودن جواب و نظائر آن ذکر می شوند.

فصل دوم بر مروری به آثار گذشته در زمینه تحلیل داده‌ها بر مبنای نرم L_1 و روشهای مرتبط با توجه خاص بر الگوریتم های نزولی تخصیص داده شده است. از آنجائیکه یک دوره عمومی دراین موضوع وجود ندارد امید می رود که این فصل یک منبع اصلی برای سایر علاقه‌مندان در نرم L_1 شود. با اشاره به بیش از ۶۰۰ مقاله و کتاب بسط نظری نرم L_1 در زمینه های مختلف آمار و اقتصادسنجی همانند استنباط آماری، خواص نمونه‌های کوچک و مجانبی، توزیع نمونه‌گیری، دستگاه معادلات همزمان و کاربرد نرم L_1 در علوم با تاکید خاص بر اقتصاد از زمان (Galilei 1632) تا حال حاضر دوره خواهد شد.

درفصل سوم الگوریتم های نزولی جدید پیشنهاد می شوند. در مطالعات قبلی

Bidabad (1987 a,b, 88a,b) الگوریتم هائی مقدماتی برای رگرسیون نرم L_1 معرفی شوند· سعی براین است که آنها را بسط داده ونوع خاصی ازروش نزولی سریع که مقدمتا" درمقالات فوق ارائه گردیدرا برای پیداکردن برآوردهای پارامترهای رگرسیون حداقل قدرمطلق خطاها بکاربندیم · بدین معنی که برای پیداکردن حداقل تابع ابژکتیونرم L_1 رگرسیون ، m-1 نقطه روی چندوجهی تابع ابژکتیوانتخاب می شوندوازاین مجموعه نقطه m باسریع ترین نزول پیدا می شود· نقطه مناسب را حذف نموده ونقطه m ام گذشته رابرای قدم نزول بعدی واردمی کنیم ·این مراحل تا رسیدن به حداقل کلی تابع ادامه می یابد·اگرچه اکثرروش های نزولی مراحل مشابهی دارند،قدم هابرای شکل خاص تابع ابژکتیونرم L_1 سازمان داده وتعدیل می شوند·دراین قسمت قضایای همگرائی جدیدی درارتباط باالگوریتم های پیشنهادی اثبات می شوندوخواص آنهاموردبحث قرارمی گیرد·توضیح قدم به قدم الگوریتم های پیشنهادی یکی ازموضوعات این فصل است ·

مقایسه الگوریتم های پیشنهادی بابعضی ازالگوریتم های موجوددرفصل چهارم انجام خواهدشد،مقایسات اساسا" برمبنای تشخیص های متفاوت توزیع های خطامی باشد·فضا و زمان اجرای لازم برای الگوریتم های منتخب درمقایسه باالگوریتم های پیشنهادی گزارش می شوند·

درفصل پنجم مسئله برآوردنرم L_1 پیوسته ملاحظه خواهدشد·باتوسعه‌روشی مشابه‌که درفصل سوم برای حالت گسسته پیشنهادشد·برآوردنرم L_1 پیوسته بدست می آید·این روش برای‌برآوردنرم های تابعی منحنی‌لورنزکه توسط Bidabad (1989 a, b) و Bidabad (1984) Gupta پیشنهادشده‌است بکاربرده خواهدشد·

فصل ششم به خلاصه ،نتیجه گیری وتوصیه برای تحقیقات بعدی درنظرگرفته شده‌است · ضمیمه برنامه‌های کامپیوتری وکتابشناسی قسمت های آخراین تحقیق می باشند·

۲- فضای نرم

درسراسراین متن نمادها وتعاریف قراردادی متعددی رابکارمی بریم ·بنابراین ،مناسب است کەنگاهی برتعاریف ،نمادها وخواص فضای نرم داشته باشیم ·اگرچه ،فضای نرم رشته طولانی درریاضیات می باشد ،این قسمت فقط بەچندصفحه جهت یادآوری محدودشده است·

باداشتن دومجموعه X و Y ،جفت مرتب x,y راکه $x\in X$ و $y\in Y$ باشدرادرنظربگیرید ·مجموعه تمام این جفت های مرتب ضرب دکارتی X و Y نامیده شده وبا $X\times Y$ آنرانشان می دهیم ·صفحه دوبعدی معمولی رامیتوان به شکل $\mathbb{R}\times\mathbb{R}$ یا $\mathbb{R}^2$ نوشت که $\mathbb{R}$ مجموعه اعدادحقیقی است ·بامجموعەهای $X_1,\ldots,X_n$ میتوان ضرب دکارتی رابه شکل زیردرنظرگرفت ·

$$\underset{i=1}{\overset{n}{\times}} X_i \tag{1}$$

اگربنویسیم $x=(x_1,\ldots,x_n)\in \underset{i=1}{\overset{n}{\times}} X_i$ ، بنابراین x_i مختصات i ام نقطه x نامیده می شود·

باهردوعضو دلخواه $u=(u_1,\ldots,u_n)$ و $w=(w_1,\ldots,w_n)$ در $\mathbb{R}^n$ قاعده ای رابرای ضرب u,w به شکل زیرمیتوانیم تعریف کنیم·

$$u.w = \sum_{i=1}^{n} u_i w_i \tag{2}$$

$u.w$ یك عددحقیقی است واگرباقاعدەفوق‌حساب شودضرب داخلی در $\mathbb{R}^n$ خوانده می شود ·بطورکلی برای هرفضای خطی دلخواه U (دریك میدان حقیقی)،ضرب داخلی بعنوان یك تابع مقدارحقیقی تعریف شده درضرب دکارتی $U\times U$ (که با $u.w$ یا $\langle u,w\rangle$ که $u\in U$ و $w\in W$ است مشخص می شود) تعریف می شودکه خواص زیررا دارا باشد·اگر $u,w,z\in U$ $\alpha,\beta\in R$ باشد،پس :

$$
\begin{aligned}
&\text{الف-۱}\quad u.w = w.u\\
&\text{الف-۲}\quad (\alpha u+\beta w).z = \alpha(u.z) + \beta(w.z)\\
&\text{الف-۳}\quad u.u \ge 0 \quad \text{و} \quad u.u = 0 \ \text{اگر وفقط فقط}\ u=0
\end{aligned}
\tag{3}
$$

یك فضای خطی بایك ضرب داخلی که دربالاتعریف شدیك فضای ضرب داخلی راتشکیل می دهند ·بەوضوح ضرب داخلی که دربالابرای $\mathbb{R}^n$ تعریف شدقواعدفوق رابرقرار

می سازدوبعنوان ضرب داخلی (اقلیدسی) معمول نامیده می شود.

اگرداشته باشیم $u=(u_1,...,u_n)$ متعلق به R^n ، فاصله بیـن u و مبداء را می توان به شکل زیرحساب نمود.

$$d(u,0)=\sqrt{\left(\sum_{i=1}^{n} u_i^2\right)}=\sqrt{(u.u)}=\sqrt{\langle u,u\rangle} \qquad (۴)$$

همینطوربرای دونقطه دلخواه $w=(w_1,...,w_n)$ و $u=(u_1,...,u_n)$ در R^n، فاصله بین u و w رامیتوان به شکل زیرحساب نمود.

$$d(u,w)=\sqrt{\left[\sum_{i=1}^{n}(u_i-w_i)^2\right]}=\sqrt{[(u-w).(u-w)]}=\sqrt{\langle (u-w),(u-w)\rangle} \qquad (۵)$$

فاصله تعریف شده فوق فاصله اقلیدسی نامیده می شود.این فاصله خواص زیررادارا ست

(۶)

ب-۱ $d(u,w)=0$ اگروفقط اگـــر $u=w$

ب-۲ $d(u,w)+d(w,z)\geq d(u,z)$

ب-۳ $d(u,w)\geq 0$ برای تمام u و w

ب-۴ $d(u,w)=d(w,u)$

باهرمجموعه دلخواه X ، اگریک تابع d $X\times X$ رابه R ببردواگـــر d شرایط بالا رابرقرارسازد، X رایک متریک می نامیم و $d(u,w)$ فاصلهبین دونقطه w,u در $X\times X$ نامیده می شود. فضای متریک با (d, X) مشخـــص می شود.

برای هرنقطه u در R^n ، فاصله اقلیدسی بین u ومبداء، $d(u,0)$ هـــم چنین نرم اقلیدسی u خوانده می شودوآنرابا $||u||$ مشخص می کنیم. بنابـــر این $d(u,w)$ رامیتوان با $||u-w||$ نشان داد. برای هرنقطه $u,w\in R^n$ و اسکالـــر $\alpha\in R$ میتوانیم خواص نرم مربوطه رابشکل زیرمشخص کنیم.

(۷)

ج - ۱ $||u||\geq 0$ و $||u||=0$ فقـــط وفقـــط اگر $u=0$

ج - ۲ $||u+w||\leq ||u||+||w||$

ج - ۳ $||\alpha u||=|\alpha|.||u||$

اگریک فضای برداری دلخواه X (نه لزوما " R^n) داشته باشیم واگرتابع مقدارحقیقی به نام نرم راتعریف کنیم که سه خاصیت فوق رابرقرارسازد، می گوئیم X یک فضـــای برداری نرم یافضای خطی نرم می باشد. بطورواضح ترهرفضای خطی نرم یک فضای متریـــک

با متریك القائی تعریف شـــده بـــــا $d(u,w)=||u-w||$ می باشد(نگاه کنیدبـــه (Takayama (1974)

مجددا " اگردونقطه $u,w \in R^n$ رادرفضای برداری X داشته باشیم متریك یعنوان تابعی از $X \times X$ به R که به نام متریك مینکوسکـــــــــی (Minkowski) معروف است به شکل زیرتعریف می شود.

$$d_p(u,w) = [\sum_{i=1}^{n} |u_i - w_i|^p]^{1/p} \qquad p \geq 1 \tag{8}$$

حال خواص ۱ـ الف تا ۴ـ الف همگی برقرارمی شوند، X یك فضای متریـك و $d_p(u,w)$ فاصله مینکوسکی دونقطه u و w در $X \times X$ خوانده‌می شود. نامساوی مثلثی ۲ـ ب برای هر $p \geq 1$ نامساوی مینکوسکی نامیده‌می شودکه همیشه برقـــرار است. اثبات رامیتوان در(1976) Vulikh و (1968) Spiegel دید. بایدتوجه داشت که نامساوی مینکوسکی برای $p<1$ برقرارنمی‌باشد. بدین ترتیـب، از آنجائیکه ۲ـ ب برقرارنمی شود $d_p(u,w)$ دیگرفاصله‌نمی تواندباشد. فضای متریـك (X,d_p) حاصل فضای L_p خوانده می شود. فضای L_p برای $p \geq 1$ یك فضـای خطی وهمچنین با نـــاخ می باشد(نگاه کنیدبه (1976) Vulikh و (1964) Akilov و Kantorovich) فاصله اقلیدسی حالت خاصی ازفاصله مینکوسکی است که درآن $P = 2$ است.

بازنقطه u رادر R^n درنظربگیرید، فاصله مینکوسکی بین u ومبداء $d_p(u,0)$ رانـــرم مینکوسکی نامیده وآنرابا $||u||_p$ مشخص می کنیم. بدین ترتیب، $d_p(u,w)$ رامیتوان به شکل $||u-w||_p$ مشخص نمود. خواص ۱ـ ج تا ۳ـ ج برای $||u||_p$ برقرارهستند. به هرحال، نرم مینکوسکی یانرم L_p رامیتوان به شکل زیرنوشت.

$$||u||_p = d_p(u,0) = [\sum_{i=1}^{n} |u_i|^p]^{1/p} \tag{9}$$

زمانیکه $p=2$ است بانرم L_2 یااقلیدسی مواجه هستیم. پی، فاصله اقلیدسی حالـت خاصی ازفاصله L_p است (نگاه کنیدبه (1985) Ralston , Rabinowitz)

روابط مهمی بین ضرب داخلی ونرم به شکل زیروجوددارد(نگاه کنیدبه(1970) Nikaido (Spiegel (1968)

۱ـ د) نامساوی کوشی ـ شوارتز

$$|\langle u,w\rangle|^2 \leq ||u||^2 . ||w||^2 \quad (۱۰)$$

وقتی تساوی خواهدبودکه فقط وفقط اگر $(u_i/w_i)=(u_j/w_j)$ برای تمام $i,j=1,...,n$.

۲ـ د) نامساوی هولدر

$$|\langle u,w\rangle| \leq ||u||_p . ||w||_q \quad (۱۱)$$

که $(1/p)+(1/q)=1, \ p,q>1$ وتساوی وقتی برقرارخواهدشدکه اگروفقط اگر $(|u_i|^{p-1}/|w_i|)=(|u_j|^{p-1}/|w_j|)$ برای تمام $i,j=1,...,n$. اگر $p=q=2$ این نامساوی به نامساوی کوشی ـ شوارتز ۱ـ د تبدیل می شود.

۳ـ د) نامساوی مینکوسکی

$$||u+w||_p \leq ||u||_p + ||w||_p \quad (۱۲)$$

که $p > 1$ تساوی وقتی برقرارمی شودکه فقط وفقط اگر $(|u_i|/|w_i|)=(|u_j|/|w_j|)$ برای تمام $i,j=1,...,n$. اگر $p=2$ باشدبه نامساوی مثلثی ۲ـ ج ساده می شود.

برای روابط ونامساوی های بیشتردرنرم نگاه کنیدبه (1987) Hramadka II et. al باتعریف فوق ازنرم درقسمت بعدموضوع راباکاربردنرم L_p یامینکوسکی در درتجزیه تحلیل های رگرسیون ادامه می دهیم.

۳ـ نرم L_p وتجزیه تحلیل های رگرسیون

دستگاه فوق تعیین معادلات زیررااداریم،

$$y = X\beta + u \quad (۱۳)$$

که y بردار $n \times 1$ متغیرهای وابسته ، X ،ماتریس $n \times m$ متغیرهای مستقل یاتوضیحی با $n > m$ ، β بردار $m \times 1$ پارامترهای مجهول و u بردار $n \times 1$ خطاهای تصادفی می باشند. مسئله پیداکردن بردارمجهول β است بطوریکه مقداربرآوردشده y نزدیک به مقدارداده شده آن باشد. یك خانواده ازروشهاکه این مقادیربرآوردشده رابدست می دهندمعیار حداقل سازی نرم L_p می باشد. (نگاه کنیدبه (1982) Narula دراین خانواده

$\|u\|_p$ برای پیداکردن بردار β حداقل می شود.

$$\min_{\beta} S = \min_{\beta} ||u||_p = \min_{\beta} ||y-X\beta||_p =$$

$$\min_{\beta} [\sum_{i=1}^{n} |y_i - x_i\beta|^p]^{1/p} = \min_{\beta} [\sum_{i=1}^{n} |y_i - \sum_{j=1}^{m} \beta_j x_{ij}|^p]^{1/p} \Longrightarrow$$

$$\min_{\beta} \sum_{i=1}^{n} |y_i - x_i\beta|^p = \min_{\beta} \sum_{i=1}^{n} |y_i - \sum_{j=1}^{m} \beta_j x_{ij}|^p \qquad (۱۴)$$

که y_i عنصر i ام y و x_i ردیف i ام ماتریس X می باشند. هر مقداری از $p\in[1,\infty]$ رامیتوان برای پیداکردن β در (۱۴) بکار برد(نگاه کنیدبه (Rice (1983) و Money et al (1978 a) ولی هرمقدار p برای توزیع خطای نوع خاصی مناسب است. نویسندگان زیادی این مسئله رابررسی کرده‌اند(نگاه کنیدبه (Barrodale (1968) و , c, 81a,b) Gonin,Money(1985a,b),Money et al(1978b,82),Barr et al(1980a , b (Sposito (1987b),Sposito,Hand,Skarpness(1983) ,Sposito,Hand(1980)

به هرحال ، توجیه p ازقضیه زیرمنتج می شود(نگاه کنیدبه (Kiountouzis (1971) و Rice, White (1964) و Hogan (1976) و (Taguchi(1974, 78)

قضیه: اگردرالگوی (۱۳) ، X غیراحتمالی و $E(u)=0,\ E(uu^T)=\sigma^2 I$ و بـا $f(u)=h.\exp(-k|u|^p)$ توزیع شده باشدکه h و k ثابت هستند و $p\in[1,\infty]$ باشد "بهترین" β باخواص حداکثراحتمال برداری است که ازحداقل کردن (۱۴) بدست می آید.

مقادیرمشخص p اهمیت بخصوصی دارند(نگاه کنیدبه (Box , Tiao (1962) و Theil (1965) و Anscombe (1967) و Thampson (1970) Zeckhauser و Sargent (1971) و Blattberg (1972) و Kadiyala (1977) و Maddala) . حداقل کردن نرم L_∞ (۱۴) بنام حداقل سازی نرم یکنواخت یاچیبیشو(Tchebyshev) یاحداقل حداکثر انحرافات خوانده می شودودارای خواص حداکثراحتمال است هرگاه که u دارای تابع توزیع احتمال یکنواخت باشد. زمانیکه p=2 است باروش حداقل مربعات مواجه خواهیم بود.دراین حالت اگرتوزیع خطاهانرمال باشدبهترین برآوردنااریب می باشد(نگاه کنیدبه(Anderson (1962) و Theil (1971)).زمانیکه p=1 است مسئله حداقل سازی نرم L_1 یانرم گرشگورین (Gershgorin) راداریم. این

مسئله همچنین حداقل یاکمترین مجموع قدرمطلق خطاها (LSAE , MSAE)،کمترین یا حداقل قدرمطلق انحرافات ،خطاها ، پسماندها یامقدارها (LAV , LAR , LAE, LAD , MAV ,MAE, MAD) برازش ،تقریب ،رگرسیون یابرآوردنرم L_1 نامیده‌می شود.

مقالات جالب Harter (1974a, b,75 a,b,c, 76) یك تاریخ شناسی تقریبا" کامل درموردبرآوردکننده‌هاکه شامل برآوردنرم L_1 نیزمی شودارائه‌می دهد. مروری درموردتحلیل های آماری برمبنای نرم L_1 توسط (Dodge (1987 انجام شده وبحثی خیلی خلاصه نیزتوسط Gentle (1977) صورت گرفته است.(1982) wellington و Narula (1987)و Narula بطورخلاصه وفشرده رگرسیون نرم L_1 رابررسی می نمایند.(Blattberg , Sargent (1971 نشان می دهندکه اگرخطاهای رگرسیون ازقانون دوم لاپلاس(توزیع نمائی دو- دم) باتابع چگالی احتمال :

$$f(u)=(1/2\theta).\exp(-|u|/\theta) \tag{۱۵}$$

تبعیت نمائیدکه $var(u)=2\theta^2$ ، حداقل کردن نرم L_1 به برآوردکننده‌حداکثر احتمال منتهی می شود.

۴- خواص برآوردکننده نرم L_1

همانندسایرمعیارها ،برآوردنرم L_1 خواص مخصوص به خودرادارامی باشدکه ازلحاظ جنبه‌های محاسباتی وآماری مهم هستند. خواص مهمترازقرارزیرمی باشند.

۴-۱- خاصیت تغییرناپذیری

یك برآوردکننده $\hat{\beta}(y,X)$ پارامترجامعه β تغییرناپذیراست اگر

$$\hat{\beta}(\theta y,X) = \theta\hat{\beta}(y,X), \qquad \theta\in[0,\infty) \tag{۱۶}$$

Koenker و Bassett (1978)و Gentle و Sposito (1976) ثابت کرده‌اندکه برآوردکننده نرم L_p زمانیکه الگوی رگرسیون خطی است تغییرناپذیراست . برآوردکننده نرم L_p برای الگوی غیرخطی عمومی تغییرناپذیرنیست خاصیت تغییرناپذیری همگنی ازدرجه یك تابع جواب $\hat{\beta}$ می باشد.

۴-۲- تبدیل متغیرها

اگر $\theta \in R^m$ باشد، تبدیل y به $y+X\theta$ مقدار بهینه $\hat{\beta}$ به اندازه θ افزایش خواهد یافت (نگاه کنید به Koenker , Bassett (1978))

$$\hat{\beta}(y+X\theta, X) = \hat{\beta}(y, X) + \theta \tag{۱۷}$$

اگر A یك ماتریس غیرمنفرد $m \times m$ باشد، با تبدیل X به XA مقدار $\hat{\beta}$ توسط پیش ضرب معکوس A تغییر می کند (نگاه کنید به Taylor (1974) و Bassett و Koenker (1978) و Koenker و Bassett (1978))

$$\hat{\beta}(y, XA) = A^{-1}\hat{\beta}(y, X) \tag{۱۸}$$

۴-۳- تحدب تابع ابژکتیو

برای نشان دادن تحدب S در (۱۴)، فرض کنید m=1 است ، تابع ابژکتیو به شکل زیر ساده می شود.

$$S = \sum_{i=1}^{n} |y_i - \beta_1 x_{i1}| = \sum_{i=1}^{n} S_i \tag{۱۹}$$

که $S_i = |y_i - \beta_1 x_{i1}|$ ، اگر S_i را بعنوان تابعی از β_1 رسم کنید، یک خط شکسته در صفحه $S x \beta_1$ خواهیم داشت. و مقدار تابع آن $\beta_{i1} = y_i / x_{i1}$ برابر صفر خواهد بود. شیب نیم خط های سمت چپ و راست β_{i1} بترتیب برابر $-|x_{i1}|$ و $|x_{i1}|$ می باشند. بنابراین ، S_i ها همگی محدب بوده و مجموع آنها S نیز محدب و در هر نقطه β_1 دارای شیبی برابر مجموع شیب های S_i ها در همان مقدار β_1 خواهد بود (نگاه کنید به Karst (1958) و Taylor (1974)).

حال (۱۴) را برای وقتی ملاحظه کنید که m=2 است ،

$$S = \sum_{i=1}^{n} |y_i - \beta_1 x_{i1} - \beta_2 x_{i2}| = \sum_{i=1}^{n} S_i \tag{۲۰}$$

که $S_i = |y_i - \beta_1 x_{i1} - \beta_2 x_{i2}|$ می باشد، میتوانیم S_i را بعنوان تابعی از β_1 و β_2 رسم کنیم. هر S_i از دو نیم صفحه در فضای $S x \beta_1 x \beta_2$ تشکیل می شود که در صفحه $\beta_1 x \beta_2$ متقاطع هستند. بدین ترتیب S_i به سمت پائین محدب است و حداقل آن برروی تقاطع دو نیم صفحه قرار می گیرد. از آنجائیکه S_i ها همگی محدب هستند، مجموع آنها رویه S نیز محدب است. بسط به m متغیر مستقل نیز به همین شکل خواهد بود. در این حالت هر S_i شامل دو نیم فوق صفحه m بعدی در فضای

$S x \beta_1 x \ldots x \beta_m$ می باشدکه‌درفوق صفحه $\beta_1 x \ldots x \beta_m$ متقاطع می باشـــد ،و همانندقبل درجهت مخالف محور S محدب است • S که درمجموع تمام این نیم فوق صفحه‌ها می باشدیك فوق رویه چندوجهی تشکیل داده‌که همچنان محدب است •

۴-۴- پسماندهای صفردرجواب‌بهینـــه

فوق صفحه رگرسیون نرم L_1 همیشه از r تا از n نقطه‌داده‌ها می گذرد ،که r مرتبـه ماتریس X می باشد •معمولا " X دارای مرتبه کامل است وبنابراین r مساوی m می باشد •بنابراین به‌تعدادپارامترهاپـسماندهای صفربرای جواب حداقل (۱۴) وجوددارد • این بدین معنی است فوق صفحه رگرسیون نرم L_1 بایداز m نقطه مشاهدات عبورکنـــد (نگاه کنیدبه (Karst 1959) و(Taylor 1974) و (Money et al (1978) و(Smith و Appa (1973) و(Kennedy (1977) و Gentle و Sposito)•

این پدیده بدلیل شکل چندوجهی S می باشد •واضح است کـه جواب حداقل بـــر روی لااقل یکی ازگوشه‌های S واقع می شودگوشه‌های S مکان هندسی تغییرات شیب های فوق رویه چندوجهی می باشدتوجه کنیداین گوشه‌هاوهمچنین لبه‌های S با لای تقاطع‌های هر زیرمجموعه m تائی از n فوق صفحه زیرمی باشند•

$$y_i - \sum_{j=1}^{m} \beta_j x_{ij} = 0 \qquad i \in \{1, \ldots, n\} \tag{۲۱}$$

ازآنجائیکه هرکدام ازاین فوق صفحه‌ها مربوط به‌یك زیرمجموعه m تائی مشخـــص از مشاهدات می باشد نتیجتا " m مشاهده وجودخواهدداشت که برروی فوق صفحه رگرسیـون قرارخواهندگرفت (نگاه کنیدبه (Taylor 1974))

۴-۵- شـرط بهینـه گـی

این شرط ازشرط لازم کون - تاکر(Tucker - Kuhn) دربرنامه‌ریزی غیرخطی بدست می آیدوتوسط (Money 1987 b) و (Gonin 1979) و Charalambous به اثبات رسیده‌است • تعریف می کنیـــــــــم

$A=\{i \mid y_i - x_i \beta^* = 0\}$ و $I=\{i \mid y_i - x_i \beta^* \neq 0\}$ ، دررگـــــــرسیــــــون

نرم L_1 خطی ،یك شرط لازم وكافی برای اینكه β^* جواب فراگیرنـــــرم L_1 باشـــد وجودضرائب $\alpha_i \in [-1,1]$ بطوری است که

$$\sum_{i \in A} \alpha_i x_i + \sum_{i \in I} sgn(y_i - x_i \beta^*) x_i = 0 \tag{۲۲}$$

(همچنین نگاه كنیدبه (1976) Dutta و Vidyasagar و El- Attar). (1973) Smith و Appa نشان دادندكه‌این جواب فوق صفحه‌ای است كـــه برای آن

$$|n^+ - n^-| \leq m \tag{۲۳}$$

كه n^+ و n^- بترتیب تعدادمشاهدات با لاوپائین فوق صفحه رگرسیون می باشد.

۴-۶- جوابهای یکتاوغیریکتا

ازآنجائیكه S یك فوق رویه چندوجهی محدب است ،همیشه دارای یك حداقل اســت . این جواب اغلب یكتاست .گاهی اوقات شكل S طوری است كه یك قسمت خط یایك چنـــد ضلعی بسته یاچندوجهی یافوق چندوجهی از S موازی فوق صفحه $\beta_1 x \ldots x \beta_m$ می شود.دراین حالت پارامترهای رگرسیون نرم L_1 یكتانیستندوبینهایت نقطه متعلق بــه خط یاچندضلعی ویاچندوجهی یافوق چندوجهی فوق جواب هستند(نگاه كنیدبه (1961) Moroney و(1973) Hartley و Sielken و(1974) Taylor و (1985) Farebrother و(1982) Sposito و(1977) Harter).

۴-۷- تجزیه تحلیل های داخلی وحساسیت

(1985) Wellington و Narula نشان دادند كه‌برآوردهای نرم L_1 متاثر ازنقاط داده‌ای بخصوصی نیستند .بدین ترتیب حذف این نقاط مقادیربرآوردشده‌پارامترهـــای رگرسیون راتغییرنمی دهد .دربحث دیگری كه آنراحساسیت برآوردهای نرم L_1 می خوانــند مقادیری راتعیین می كنندكه مقدارمتغیرپاسخ y_i میتوانـندبدون اینكه بربرآوردپارامتر ها تاثیربگذاردتغییركند .مشخصا " ازمقدار y_i كاهش یاافزایش یابدبدون اینكه‌علامت u_i راتغییردهد ،جواب پارامترهاتغییرنخواهدكرد(نگاه كنیدبه (1809) Gauss

(1987) Farebrother).

برای توپولوژی تقریب نرم L_1 وخواص آن نگاه کنیدبه (1965) Rivlin و Kripke (1987 و Vajda (1987 و Hromadka II et al (1987 . سایر خواص رگرسیون نرم L_1 توسط (1976,77) Sposito و Kennedy Gentle (1977) Assouad و (1980) Kennedy , Gentle Sposito و (1987,88a, b) Bassett بحث می شود.

فصل دوم

" مروری بر مطالعات پیشین در تحلیل های آماری در نرم L_1 وروشهای مربوطه "

۱- مقـــــدمـــــه

اگرچه نرم L_1 یك مبحث قدیمی درعلم است ،امانبودن یك مقاله یاکتاب عمومی در این زمینه سبب شدتایك صورت نسبتا " کامل ازمنابع رادراین فصل جمع آوری نمائیم۰روشهای مرتبط بانرم L_1 خیلی گسترده هستندوتلخیص آنان کاری دشوارمی باشد۰ بههرحال،سعی شدهاست که نگاهی گذرابرتقریبا " تمام زمینههای مرتبط داشته باشیم ۰موضوعات به صورت قسمت های مجزاطرح شدهاندتابنابراین خواننده بتواندبدون ازدست دادن پیوستگی موضوعاز قسمت هائی صرف نظرنمائید۰اگرچه بعضی ازموضوعات دراین فصل کمتردرارتباط باموضوع فصول بعدقرارمی گیرد ،بدلیل نبودن یك بررسی کامل برنوشتجات پیشین دراین زمینه ،آنها راحذف نکردیم وامیدمی رودکه مورداستفاده علاقهمندان این رشته قرارگیرد۰

ساختاراین فصل ازقرارزیراست ۰ تاریخ شماری وتوسعه تاریخی نرم L_1 دردوره ۱۶۳۲ تا ۱۹۲۸ موضوع قسمت بعدی است ۰ مقالات اصلی اززمان گالیلهذکرمی شوند۰الگوریتمهای ابداع شدهدراین دوره آنقدرسازمان یافته نیستندکه بتوانندرگرسیون نرم L_1 رادرحالت عمومی حل نمایند۰درقسمت بعدالگوریتم های محاسباتی که متعلق به زمان بعداز ۱۹۲۸ می باشدرابا شیوهای دیگرموردبحث قرارخواهیم داد۰این دوره ازسال ۱۹۲۸ تابه حال زمان الگوریتم های متعالی تلقی می شودکه درسه گروه روشهای نزولی ،مستقیم ،نوع سیمپلکس وسایرطبقه بندی می شوند۰ازآنجائیکه تاکیدفصول بعدی برمبنای پیشنهادروش های نزولی جدیدمی باشد،قسمت ۳ـ۱ بیشترتشریحی می باشد۰این قسمت سعی برتلخیص روشهای نزولی مستقیم درارائه مراحل عملیاتی آنان دارد،الگوریتم هائی که براساس روش سیمپلکس برنامهریزی خطی قراردارنددرقسمت ۳ـ۲ بحث خواهندشد۰قسمت بعدبهمروری برالگوریتم هائی که متعلق به دوقسمت فوق نمی باشندتعلق دارد۰مسئله مقداراولیه برای شروع الگوریتم های مختلف درقسمت ۳ـ۴ بحث خواهدشد۰قسمت ۳ـ۵ به برنامههای کامپیوتری موجوداشاره می نماید۰قسمت ۳ـ۶ مطالعات تطبیقی محققین قبلی درکارآئی نسبی الگوریتم های موجودرابررسی می نماید۰روشهای محاسباتی رگرسیون غیرخطینرم L_1 و اشاره بهحالت نرم L_p موضوع قسمت های بعدی می باشد۰ درقسمت ۴ کاربردنرم L_1 دردستگاه معادلات همزمان رابررسی می نمائیم ۰درقسمت بعدجنبههای آماری نرم

L_1 بطورخلاصه ذکرگردیده‌ونگاهی به توزیع نمونه‌گیری ،استنباط آماری،آمارچند متغیره ،برآوردچگالی غیرپارامتری وآمارنیرومندنرم L_1 خواهیم داشت ۰ کاربرد نرم L_1 باتاکیدبیشتردرمسائل اقتصادی موضوع قسمت ۳ـ۶ می باشد۰درقسمت ۳ ـ ۷ سایرمشتقات نرم L_1 بطورخلاصه ذکرخواهندشد۰

۲ـ تاریخ شماری وتوسعه تاریخی (۱۹۲۸ ـ ۱۶۳۲)

مبداء برآوردنرم L_1 رامیتوان به (1632) Galilei نسبت داد ۰ در تعیین موقعیت یك ستاره تازه کشف شده ،وی حداقل اصلاح ممکن رابرای حصول نتیجه‌قابل اطمینان پیشنهادمی نماید(نگاه کنیدبه (1987) Ronchetti برای چند نقل وقول مستقیم)۰ (1757) Boscovich برای اولین بارحداقل مجموع قدر مطلق خطاهارافرموله کرده‌ودرجهت برازاندن بهترین خط برای سه جفت مشاهده داده شده‌یا بیشتربرای الگوی رگرسیون دومتغیره ساده آنرابکارمی برد۰وی خط رگرسیون رامقیدبه‌عبور ازمیانگین های نقاط مشاهده می کند ،بدین معنی که،

$$\begin{aligned} &\min_{\beta_0,\beta_1} : \sum_{i=1}^{n} |y_i-\beta_0-\beta_1 x_{i1}| \\ &\text{s.to:} \sum_{i=1}^{n} (y_i-\beta_0-\beta_1 x_{i1})=0 \end{aligned} \tag{1}$$

(1760) Boscovich یك راه حل هندسی ساده برای پیشنهادقبلی خودارائه می نماید۰این مقاله توسط (1961) Eisenhar و(1973) Sheynin بحث شده‌است ۰دریك نامه Boscovich مسئله رابرای Simposon طرح می نمایدو Simpson یك راه حل تحلیلی برای مسئله ارائه می دهد(نگاه‌کنیدبه (1984) Stigler)۰

(1773) Laplace یك فرمولاسیون جبری یك الگوریتم برای رگرسیون نرم L_1 که‌ازمرکزمشاهدات عبورمی کندرامهیا می سازد۰در (1779) Laplace بسط رگرسیون نرم L_1 به مشاهدات باوزن های مختلف نیزبحث شده است ۰(1804)

Prony (1779) Laplace یك تعبیرهندسی ازروش ارائه نموده‌وآنرابا سایر روش ها ازطریق یك مثال مقایسه‌می نماید. Svanberg (1805) روش Laplace رادرتعیین یك مدارنصف النهاربکارمی بنددو(1809)

Von Lindenau این روش رابرای تعیین‌مدارنصف النهاربیضوی استفاده‌می نماید.

(1809) Gauss حداقل کردن مجموع قدرمطلق خطاها رابدون قیدپیشنهاد می کندوی نتیجه می گیردکه این معیارلزوما" m تا ازپسماندهارا مساوی صفرقرارمی دهد که m برابرتعدادپارامترها می باشدوهمچنین جواب حاصل ازاین روش اگرمقدارمتغیر تابع افزایش یاکاهش یابدبدون اینکه علامت پسماندمربوطه‌را تغییر ندهدتفاوتی نخواهدکرد. این نتیجه گیری اخیرا" توسط (1985) Wellington و Narula مورد بحث‌واقع شدکه درفصل قبل تحت عنوان تجزیه تحلیل های درونی وحساسیت موردبحث قرار گرفت. وی همچنین اشاره می نمایدکه برآوردکننده‌های Boscovich یا Laplace که‌مجموع قدرمطلق خطاهارابا قیدصفربودن مجموع پسماندهاحداقل می نماید. لزوما" $m-1$ پسماندرامساوی صفرقرارمی دهد.(نگاه کنیدبه (1981) Stigler و(1987b) Farebrother).

(1816) Mathieu روش لاپلاس رابرای محاسبه دوری ازمرکززمین استفاده می کند.(1916) Van Beeck - Calkoen ازاستفاده معیارحداقل قدرمطلق هادربرازاندن معادلات منحنی الخط حاصل ازتوانهای متغیرمستقل پشتیبانی می نماید.

(1818) Laplace معیار Boscovich رادوباره‌مورداستفاده‌قرارداده ویك راه‌حل جبری ارائه می نماید(نگاه کنیدبه (1987b) Farebrother).

اگر $\bar{x}_1$ و $\bar{y}$ میانگین‌های x_{i1} و y_i باشند،بنابراین،

$$\beta_0 = \bar{y} - \beta_1 \bar{x}_1 \qquad (۲)$$

مقدار β_1 بااستفاده ازرابطه زیرپیدامی شود.

$$\min_{\beta_1}: S = \sum_{i=1}^{n} |\tilde{y}_i - \beta_1 \tilde{x}_{i1}| \qquad (۳)$$

که $\tilde{x}_{i1}$ و $\tilde{y}_i$ انحرافات x_{i1} و y_i ازمیانگین هریك می باشد. با

جابجائی مشاهدات بەترتیب نزولی مقادیر $y_i^{\sim}/x_{i1}^{\sim}$ ، لاپلاس اشاره می نمایدکه S بینهایت است اگر β_1 بینهایت باشدومادامیکه β_1 کاهش می یابد S نیزکم می شود٠زمانیکه β_1 بەمقداربحرانی $y_t^{\sim}/x_{t1}^{\sim}$ می رسدمقدار S مجددا " شروع بەافزایش می نماید٠این مقداربحرانی β_1 زمانی اتفاق می افتد که

$$\sum_{i=1}^{t-1}|x_{i1}^{\sim}| < \tfrac{1}{2}\sum_{i=1}^{n}|x_{i1}^{\sim}| \leq \sum_{i=1}^{t}|x_{i1}^{\sim}| \qquad (۴)$$

این روش برای پیداکردن مقدار β_1 میانه وزنی نامیدەمی شودوبعدا " درالگوریتم های بسیاردیگری مانند(1930) Rhodes ، (1940) Singleton ،(1958) Karst ،(1980) Steiger و Bloomfield (88 a , b) Bidabad (1987 a,b) موردا ستفادەقرارمی گیرد٠(1987a, b , 88a , b) Bidabad شرط (۴) راازطریق مشتق گسسته بدست می آورد٠

(1824) Fourier رگرسیون حداقل قدرمطلق پسماندهارابەطریق روشی کەالان آنرابرنامەریزی خطی میشناسیم فرمولەمی نماید،یاحداقل کردن تابع ابژکتیوخطی باتوجەبه قیدهای نامعادله خطی ٠

(1883) Edgeworth یك بحث فلسفی دراختلافات مابین حداقل کردن میانگین مربع خطاهاومیانگین قدرمطلق خطاهاارائه می نماید٠در(1887a, b) Edgeworth یك روش ساده برای انتخاب پارامترهای رگرسیون پیشنهادمی شود٠با ثابت نگاەداشتن m-1 پارامتر،وی روش لاپلاس رابرای تعیین مقداربهینه پارامتر دیگربکارمی بندد٠باتکراراین عمل برای دامنەای ازمقادیر m-1 پارامترثابت نگهداشته شده ،مجموعەای ازنتایج برای m انتخاب ممکن پارامترآزادبدست می دهد Edgeworth محدودیت عبورازمرکزدادەهارالغومی نماید٠(1887) Turner مسئله جوابهای غیریکتارابامعیارحداقل قدرمطلق خطادرقالب گونه تصویری روش (1887a) Edgeworth بعنوان نارسائی موجودروش وی مورد بحث قرارمی دهد٠ (1888) Edgeworth بحث Turner راباپیشنهادیك روش دوم برای رگرسیون حداقل قدرمطلق خطای خطی سادەکەدرآن مکان هندسی میانەهای روش اولش مورداستفاده واقع نمی شودپاسخ می دهد٠دراین مقاله Edgeworth پیشنهادتحلیل گرافیکی

Turner برای مراحل لازم به رسیدن به‌جواب حداقل را استفاده می نماید.

قبل از اشاره به روش "دومیانه" (1923) Edgeworth ،باید خاطرنشان کرد که (1902) Bowley مقاله (1902) Edgeworth را باگونه‌ای از روش دومیانه که بعدا" توسط (1923) Edgeworth ارائه می شود تکمیل می نماید. روش وی انتصاب وزن ها را به خطاها نادیده می گیرد.

(1923) Edgeworth مسئله عمومی تربرآورد پارامترهای رگرسیون خطی ساده را با حداقل کردن مجموع وزنی قدرمطلق پسماندها موردبحث قرار می دهد. وی دلائل منطقی روش را مطرح کرده و توسط چندمثال کاربرد آن را توجیه می نماید. وی همچنین جواب غیریکتا را مدنظرقرار می دهد. روش پیشنهادی وی روش "دومیانه" نامیده می شود.

(1926-28) Estienne جایگزینی نظریه کلاسیك خطاهای داده‌ها برمبنای حداقل مربعات با چیزی که وی نظریه منطقی براساس روش حداقل قدرمطلق خطا می نامد را پیشنهاد می نماید. (1928) Bowely نوشتجات Edgeworth درآمار ریاضی را خلاصه می نماید که شامل کارهای وی درزمینه رگرسیون نرم L_1 نیز می باشد. (1928)

Dufton نیز یك روش گرافیکی برای برازاندن خط رگرسیون پیشنهاد می نماید.

(1987 b) Farebrother مقالات مهم در رابطه با رگرسیون نرم L_1 را برای دوره ۱۹۳۰-۱۷۹۳ خلاصه می کند. برای منابع بیشتر به (1969) Crocker و Dielman (1984) و Harter (1974a, b, 75a, b, c, 76) رجوع نمائید.

تا سال ۱۹۲۸، تمام الگوریتم ها برای رگرسیون خطی ساده پیشنهاد شده بود. گرچه بعضی از آنان دارای بیان جبری هستند، ولی آنچنان سازمان یافته نیستند که بتوانند مسئله رگرسیون چندمتغیره نرم L_1 را حل نمایند. درقسمت بعد روش های محاسباتی پیشرفته تر برای رگرسیون نرم L_1 ساده و چندمتغیره را به دلیل وجود رجعت هائی به صورت شمارش غیر تاریخی مقالات موردبحث قرارخواهیم داد. این دوره از سال ۱۹۲۸ تاکنون را زمان الگوریتمهای مدرن درموضوع نرم L_1 می توان محسوب نمود.

۳ـ الگوریتم هـای محـاسبـاتـی

اگرچه فرم بسته جواب رگرسیون نرم L_1 تاکنون پیدا نشده ،الگوریتم های زیـادی برای حداقل کردن تابع ابژکتیوآ ن پیشنهادشده است (نگاه کنیدبه ، Cheney (1966) و Chambers (1977) و Dielman, Pfaffenberger (1982, 84))

بەطورکلی میتوان تمام الگوریتم های نرم L_1 رادرسه‌گروه اصلی زیرطبقه‌بندی نمود.

الف) الگوریتم های نزولی مستقیم

ب) الگوریتم های نوع سیمپلکس

ج) سایرالگوریتم هـا

که متعاقبا" درقسمت های زیربررسی خواهندشد.

۳ـ۱ـ الگوریتم هـای نزولی مستقیم

الگوریتم هائی که دراین طبقه قرارمی گیرنداساسا" به دنبال یك مسیرپرشیب هستندکه برروی چندوجهی تابع ابژکتیورگرسیون نرم L_1 نزول نمایند. اگرچه روش لاپلاس (کـه قبلا" توضیح داده شد) نوع خاصی ازالگوریتم های نزولی مستقیم است ،مبداء استفاده‌ازاین روش درزمینه‌نرم L_1 رامی توان به‌الگوریتم های اجورث که درقسمت قبل بیان گردیدنسبت داد.

Rhodes (1930) بدلیل کاربربودن راه حل گرافیکی اجورث روش دیگری رابرای الگوی عمومی خطی پیشنهادمی نمایدکه به شکل زیرخلاصه می گردد(نگاه کنیـدبـه Farebrother (1987b) فرض کنید n معادله با $m < n$ پارامتر نامعلوم داریم. برای پیداکردن جواب نرم L_1 این "دستگاه معادلات فوق تعیین" قدمهای زیرپیشنهادمی شود.

قدم اول) $m-1$. معادله را بطـور دلخواه انتخاب می کنیم.

قدم دوم) این معادلات رابرای $m-1$ پارامترحل می نمائیم.

قدم سوم) پارامتر m ام باقی مانده رابا روش لاپلاس بدست می آوریم.

قدم چهارم) معادله حاصل درقدم سوم راشناسائی نموده وآنرابه $m-1$ معادله قـدم اول اضافه‌می کنیم.

قدم پنجم) اگرمجموعه m معادله m بارتکرارشدتوقف می کنیم ،ودرغیراینصـــورت قدیمی ترین معادله رادورمی اندازیم وبه قدم دوم می رویم.

Rhodes (1930) الگوریتم خودرابایك مثال توضیح می دهدوهیچگونه برهـانـی برای‌همگرائی آن ارائه‌نمی نماید. Bruen (1938) روشهای رگرسیون نـــرم L_1 متقدمین خودرامرورمی نماید.وی همچنین رگرسیون نرم های L_1 و L_2 و L_∞ را مقایسه می نماید.

Singleton (1940) روش نزول سریع‌کوشی (نگاه کنیدبـــه Paink (1974)) رابرای رگرسیون نرم L_1 درحالت عمومی بکارمی برد. دراین مقاله یــك تعبیرهندسی ازگرادیان برچندوجهی نرم L_1 وقضایایی درباره وجودویکتابودن جـواب وخاصیت تحدب ارائه شده‌است ،اگرچه این مقاله‌خیلی روشن نوشته شده‌است ،قدم های زیـر الگوریتم وی راخلاصه می نماید.

قدم اول)نقطه دلخواه $\beta_j{(0)},\ j=1,\ldots,m$ راانتخاب کنید.

قدم دوم) گرا یان زیرراتعیین نمائید،

$$g_j{(0)}=-\sum_{i=1}^{n}\operatorname{sgn}(u_i \hat{}\,|\,\beta_j{(0)},j=1,\ldots,m)x_{ij}$$

قدم سوم) این مقادیرراحساب کنید

$$w_i{(0)}=\sum_{j=1}^{m}x_{ij}g_j{(0)},\quad z_i{(0)}=y_i-\sum_{j=1}^{m}x_{ij}\beta_j{(0)}$$

قدم چهارم) مقدار $t(0)$ رابعنوان میانه وزنی عبارت $\sum_{i=1}^{n}|w_i t-z_i|$ باروش لاپلاس بدست آورید.مقدار t طول حرکت درجهت گرا یان می باشد.

قدم پنجم) حساب کنید $\beta_j{(1)}=\beta_j{(0)}+g_j{(0)}t(0)$

قدم ششم) شرط بهینه بودن راآزمایش کنید. Singleton یك شرط برای توقف می‌دهد ولی کاملا" واضح نیست .دراین جاهرشرط مناسب دیگری برای تابع‌نـرم L_1 را میتوان جایگزین نمود.

قدم هفتم) این قدم به‌خوبی توسط Singleton توضیح داده‌شده‌است .دراین مرحله وی سعی برانتخاب بهترین گرادیان ازمیان گرادیان های قابل استفاده‌دارد .بدون وجود این قدم الگوریتم وی هنوزعملی می باشد،زیرا ،قدم های فوق همه‌همان قدم هـای روش نزول سریع‌کوشی می باشدوبه‌جای انتخاب بهترین گرادیان میتوانیـــم الگوریتم رابارفتن به قدم دوم ادامه دهیم.

(1956, 57) Bejar برملاحظه پسماندهاتابردارپارامترهاتاکید می نماید. وی یك الگوریتم با ماهیت (1930) Rhodes رانیزپیشنهاد می نماید. به هرحال الگوهای موردنظروی خطی ودارای دووسه پارامترمی باشند.

(1958) Karst یك مقاله تشریحی برای الگوهای رگرسیون یك ودوپارامتری تنظیم می نماید. دراین مقاله، وی بدون اشاره به کارهای انجام شده قبلی، عملا "بەپیشنهاد لاپلاس برای حل الگوی مقیدیك پارامتری می رسدوبرای الگوی دوپارامتری یك الگوریتمشبیه به (1930) Rhodes راپیشنهادمی نماید. روش برخوردوی بەمسئله همهندسی و هم جبری می باشدوهیچ برهانی برای همگرائی روش ترجیحی خودارائەنمی نماید. (1974) Sadovski بااستفاده ازروش " مرتب کردن حبابی " ساده الگوریتم Karst رابافرترن بەاجرادرمی آورد. (1976) Sposito خاطرنشان ساخت که برنامه Sadovski همیشه همگرانیست. (1976) Sposito , Smith (1987c) Farebrother الگوریتم دیگری رابرای رفع این مسئله پیشنهادمی نمایند. برنامه Sadovski رابەزبان پاسکال نوشته وبابکارگیری روش " مرتب کردن درج مستقیم" آنرابهبودمی بخشد.

(1967 b) Usow الگوریتمی برای تقریب نرم L_1 دادەهای گسستەپیشنهاد نموده واثبات می کندکەاین الگوریتم درقدم هائی باتعدادمحدودهمگرامی باشد. یك الگوریتم مشابەبرای تقریب نرم L_1 دادەهای پیوسته در (1967a) Usow بیان گردیدەاست.

اگر $f(x)$ ازیك فاصله برخط راست حقیقی تعریف شودوهمچنین اگرتوابع پیوسته مستقل خطی $\Phi_1(x),\ldots,\Phi_m(x)$ برزیرمجموعه محدود $X=\{x_1,\ldots,x_n\}$ و $m<n$ راداشته باشیم تابع " چندجملەای" $L(\beta,x)=\sum_{j=1}^{m}\beta_j\Phi_j(x)$ رادرنظر بگیرید. در (1967b) Usow تابع زیرحداقل می شود:

$$\min_{\beta}: \; S(\beta)=\sum_{i=1}^{n}|L(\beta,x_i)-f(x_i)|=\sum_{i=1}^{n}\left|\sum_{j=1}^{m}\beta_j\Phi_j(x_i)-f(x_i)\right| \quad (۵)$$

که β برداری باندازه m می باشد. فرم فوق کلی است، اگرقراردهیم $f(x_i)\equiv y_i$ و $\Phi_j(x_i)\equiv x_{ij}$ تابع S تبدیل به تابع ابژکتیورگرسیون خطی می شود. حال مجموعه K رابه شکل زیرتعریف می کنیم:

$$K=\{(\beta,d)\,|\,(\beta,d)\in E^{m+1},\, S\leq d\} \quad (۶)$$

بنابراین مجموعه K یك چندوجهی محدب است که تارکهای آن وقتی m یابیشتر ازنقاط X مقدار $L(\beta, x)-f(x)$ راصفرکنندایجادمی شوند. الگوریتم $Usow$ نزول کردن برروی K ازیك تارك به تارك دیگردرا متدادنبش های پیوسته پلی تاپ می باشدبطوریکه تارك های واسطه‌ای معینی حذف می شوند. این نزول تارسیدن به پائین ترین تارك (β^*, d^*) ادامه می یابد. برای روشن ترکردن الگوریتم فرض کنیدکه بر روی تارك (β^k, d^k) در K قرارداریم وچندجمله‌ای ($L(\beta^k, x)$) m نقطه‌از X بنام $U^k=\{u_1{}^k, \ldots, u_m{}^k\}$ رادرونیابی می کندبدین ترتیب ،

$$L(F^k, x)=\sum_{i=1}^{m} f(u_i{}^k)\pi_i(x) \tag{۷}$$

که‌درآن

$$F^k=(f(u_1{}^k), \ldots, f(u_m{}^k)) \tag{۸}$$

و

$$\pi_i(x)=\sum_{j=1}^{m} a_{ji}\Phi_j(x) \quad i=1, \ldots, m \tag{۹}$$

mضریب a_{ji} به‌شکل زیرمحاسبه می شوند. ماتریس $(\Phi_j(u_i{}^k))$ رابرای $i, j=1, \ldots, m$ تشکیل می دهیم. دوبردار $m\times 1$ برای هر $\pi(x)$ و $\Phi(x)$ $x\in X$ که‌عناصرآن بترتیب $\pi_i(x)$ و $\Phi_i(x)$ برای $i=1, \ldots, m$ می باشدتعریف می شوند. پس میتوانیم بدست آوریم،

$$\pi(x)=[(\Phi_j(u_i{}^k))^T]^{-1}\Phi(x) \tag{۱۰}$$

a_{ji} هاعناصرردیف i ام‌ماتریس $[(\Phi_j(u_i{}^k))^T]^{-1}$ هستندکه با لانویس T معرف ترانسپوزمی باشد. اگر e_i یك بردارصفربه‌اندازه m باشدکه‌عنصر i ام آن‌مساوی یك هست ،بنابراین اگربرای بعضی مقادیر δ ، $S(F^k-\delta e_i)<S(F^k)$ یك T_j وجوددارد‌بطوریکه $T_j\delta>0$ و $S(F^k-T_je_i)<S(F^k)$ همچنین

$$S(F^k - T_je_i) = \min_t \{S(F^k - te_i)\} \tag{۱۱}$$

و $((F^k-T_je_i), S(F^k-T_je_i))$ یك تارك است. به عبارت دیگر

نقطه $u_i{}^k \in U^k$ رامیتوان بانقطه $u_i{}^{k+1} \in \{X-U^k\}$ تعویض نمـــود بطوریکه چندجملهای $L(\beta^{ki}, x)$ $U_i{}^k=\{u_1{}^k, \ldots, u_i{}^{k+1}, \ldots, u_m{}^k\}$ رادرونیابی نمایدو $S(\beta^{ki}) < S(\beta^k)$. $S(\beta^{ki})$ حداقل تمام نرم های حاصل است اگر $u_i{}^k$ بانقاط مختلف مجموعه $\{X-U^k\}$ همانطورکه رابطه با لانشان می دهـــد جایگزین شدهباشد .

بارفتن ازتارك $(\beta^k, S(\beta^k))$ به $(\beta^{ki}, S(\beta^{ki}))$ یك تارك یابیشتردر K را میتوان حذف نمود .نزدیکترین تارك به $(F^k, S(F^k))$ وزیرآن روی لبه موازی بامحور مختصات فضای پارامتر i ،یابه عبارت دیگرتارك $((F^k-t_r e_i), S(F^k-t_r e_i))$ بهشکل ز یربدست می آید .

$$|t_r| = \min_s \left\{\frac{|L(F^k, x_s)-f(x_s)|}{|\pi_i(x_s)|}\right\} \qquad x_s \in \{X-U^k\} \qquad (۱۲)$$

نقطه x_r بهشکل زیرمشخص می شود ،

$$sgn[L(F^k, x_l) - f(x_l)] = sgn[L(F^{ki}, x_l) - f(x_l)], \quad x_l \in \{X-U^k-x_r\} \qquad (۱۳)$$

حال ،اگریك δ وجودنداشته باشدكه $S(F^k-\delta e_i) < S(F^k)$ باشـد ، $S(F^k)$ نمی توانست باحرکت در K برروی لبه موازی بامحورمختصات فضای پارامتر i ام کاهش یابدو $u_i{}^k$ نباید بانقطهدیگری ازمجموعه $\{X-U^k\}$ جایگزین شود . این ترجیـــع m بارتکرارمی شود ،به ترتیب هربار برای هرنقطه در U^k .کل مراحل مجددا " به تعدادمحدودی تکرارمی شودتاجواب (β^*, d^*) بدست آید(نگاه کنیدبه(1974 Abdelmalek)

رابطه این الگوریتم باروش سیمپلکس توسط (1974) Abdelmalek بحـث می شود .وی نشان می دهدکه الگوریتم Usow کاملا" معادل الگوریتم ثانویه سیمپلکـــس است که برای یك الگوی برنامهریزی خطی با متغیرهای محصورغیرمنفی بکاربردهشودویـــك دورتکراردراولی معادل یك دوریابیشتردردومی می باشد . (1980) Steiger و Bloomfield یك الگوریتم کارابرمبنای پیشنهاد Usow کهدربالاذکرگردیدابداع می نمایند .

(1971) Sharpe بابکارگیری رگرسیون نرم L_1 درسهام ونرخ بازگشت آن ،الگوریتمی برای الگو ی رگرسیون خطی دوپارامتری ارائه می نماید٠وی بحث می نماید که برای الگوی ساده باتابع ابژکتیو ،

$$S = \sum_{i=1}^{n} |y_i - (\beta_0 + \beta_1 x_{i1})| \tag{۱۴}$$

بایدبتوانیم نیمی ازنقاط رابه بالاونیمی دیگررابه‌پایین خط رگرسیون منصوب نمائیم٠بازاء هرمقدارمعین از $\beta^{\sim}_1$ ،میتوانیم β_0 رابعنوان میانه $\beta_{0i}=y_i-\beta^{\sim}_1 x_{i1}$ بدست آوریم ٠حال نقاط راطوری جدامی کنیم که :

$$S=\sum_{i_{above}}^{n_{above}} [y_i-(\beta_0+\beta_1 x_{i1})] - \sum_{i_{below}}^{n_{below}} [y_i-(\beta_0+\beta_1 x_{i1})] \tag{۱۵}$$

مجددا " جملات رامرتب نموده وچون $n_{above}=n_{below}$ داریم ،

$$S = k_1 + k_2\beta_1 \tag{۱۶}$$

که

$$k_1 = \sum_{i_{above}}^{n_{above}} y_i - \sum_{i_{below}}^{n_{below}} y_i \tag{۱۷}$$

$$k_2 = -\sum_{i_{above}}^{n_{above}} x_{i1} + \sum_{i_{below}}^{n_{below}} x_{i1} \tag{۱۸}$$

کل روش حل رامیتوان بدین شکل خلاصه کرد٠هرمقداری برای β_1 رامی توان درابتـــــدا انتخاب کردهبامحاسبه β_{0i} نقاط بالا وپائین خط راجدامی کنیم ٠بااستفـــــــادهاز معادلات (۱۶) تا (۱۸) قسمت مربوط به‌رابطه S و β_1 محاسبه‌می شود،علامت K_2 بیانگرجهت مناسب برای دوربعدمیباشد٠اگر K_2 مثبت است ،مقادیرکمتر β_1 فقط باید ملاحظه شوند٠اگر K_2 منفی است ،فقط مقادیربیشتر β_1 بایدملاحظه شوند٠اگر K_2 مساوی صفراست ،مقداراولیه β_1 جواب است ٠

زمانی که یك خط مرزی (میانه β_{0i}) وجهت جستجومعین شد،نزدیکترین تقاطـــــع خط مرزی فعلی باخطوط دیگرراپیدامی کنیم ٠محاسبات رابامقایسه شیب ها(مقادیر x_{i1}) برای تعیین اینکه دوخط دردامنه موردنظرمتقاطع‌هستندیانه‌می توان کاهش داد(بنابراین از عملیات تقسیم اضافه اجتناب می شود)مقادیرجدید K_1 و K_2 راحساب می کنیم ٠اگراین عملیات سبب شدکه علامت K_2 تغییرنماید،جواب بدست آمده‌است والگوریتم متوقـــف می شود٠

Sharpe (1972) Srinivasan و Rao روش Sharpe را بعنوان جواب فرمولاسیون این مسئله به شکل برنامه‌ریزی خطی ثانویه پارامتری تفسیر می نمایند. آنها یک روش متفاوت ولی تقریبا " ازلحاظ کارآئی مشابه برای حل همان مسئله ارائه می نمایند. (1980) Brown روشی متمایزولی مشابه با (1923) Edgeworth و (1971) Sharpe طرح می نماید. وی برخواص میانه برآورد کننده تاکیدمی نماید. مباحث این سه روش دردیدگاه گرافیکی طراحان آنان می باشد.

(1979) Kawara نیزیک روش گرافیکی برای الگوی رگرسیون ساده رابسط می دهد.

(1978) Sinclair و Conn و Bartels روش (1976) Conn رابرای حل نرم L_1 دستگاه خطی فوق تعیین استفاده می نمایند. دیدگاه آنان روش حداقل کردن توابع مشتق پذیرتیکه‌ای می باشد. این الگوریتم رامیتوان به شکل زیرخلاصه نمود.

قدم صفر) مقداردلخواهی برای β درنظرمی‌گیریم.

قدم اول) الف) مجموعه اندیس $I = \{i_1, \ldots, i_m\} = \{i \mid x_i^T\beta - y_i = 0\}$ رامشخص می کنیم فرض می کنیم ردیف I ام X برابراست با $X_I = [x_{i_1}, \ldots, x_{i_m}]$ که $i_j \in I$ وفضای تهی $N = N(X_I^T) = \{\delta \mid x_i\delta = 0, i \in I\}$ پروژکتورعمودبر روی N یا P_N مشخص می شود.

ب) I^c رابعنوان مکمل مجموعه I درنظرمی‌گیرم. بردار $h = \sum_{i \in I^c} \operatorname{sgn}(x_i^T\beta - y_i)x_i$ رامحاسبه‌می نمائیم.

ج)مقدار $p = -P_N h$ که تصویر $-h$ برفضای تهی X_I است مادامیکه‌این تصویرغیرصفراست رامحاسبه‌میکنیم. اگر $p \neq 0$ بگذارید $g = h$ وبه‌قدم‌دوم بروید.

د) مقدار w رابرمبنای $h = X_I w = \sum_{j=1}^{m} w_j x_{i_j}, i_j \in I$ محاسبه می نمائیم.

ه) اگر $|w_j| \le 1$ برای $j = 1, \ldots, m$ توقف می کنیم. در این حالت β بهینه می‌باشد.

و) مقدار $i_{j_0} \in I$ راطوری پیداکنیدکه $|w_{j_0}| > 1$.

ز) مجموعه I را به $I-\{i_{j0}\}$ تبدیل نموده و تغییرات مربوطه را در X_I و N انجام می دهیم، مقدار $p=-sgn(w_{j0})P_N x_{i_{j0}}$ را محاسبه نموده و $g=h-sgn(w_{j0})x_{i_{j0}}$ را قرار می دهیم.

قدم دوم) عناصر $A=\{\alpha_l \mid l\in I^c, \alpha_l=(x_l{}^T\beta-y_l)/x_l{}^T p, \alpha_l>0\}$ را تعیین نموده و طوری آنها را مرتب می کنیم که $0<\alpha_{l_1}<\alpha_{l_2}<...<\alpha_{l_t}$. قرار می دهیم $\tau=1$.

قدم سوم) اگر $p^T g\geq 2sgn(x_{l_\tau}{}^T\beta-y_{l_\tau})p^T x_{l_\tau}$ به قدم پنجم می‌رویم.

قدم چهارم) مقدار g را به $g-2sgn(x_{l_\tau}\beta-y_{l_\tau})x_{l_\tau}$ و τ را به $\tau+1$ تغییر داده و به قدم سوم می رویم.

قدم پنجم) β را به $\beta+\alpha_{l_\tau}p$ تغییر داده و به قدم اول می رویم.

این الگوریتم همچنین برای حالت دیژنرسی نیز تکمیل شده است (نگاه کنید به (1976) Sinclair و Conn و Bartels)، Conn و Bartels (1977) نشان می دهند که چگونه رگرسیون های نرم L_1، نرم L_1 مقید، نرم L_∞ و برنامه‌ریزی خطی همگی میتوانند بعنوان یك مسئله حداقل سازی خطی تیکه‌ای بیان شوند، فرض کنید U و v به‌ترتیب به اندازه‌های $p\times m$ و $1\times p$ باشند. تابع زیر را ملاحظه کنید.

$$\Phi(\beta) = h^T\beta + \Sigma_i |y_i-x_i{}^T\beta| + \Sigma_j \max(0, v_j-u_j{}^T\beta) \tag{۱۹}$$

که $y_i-x_i{}^T\beta$ نشانگر عنصر i ام بردار پسماند $y-X\beta$ و $v_j-u_j{}^T\beta$ معرف عنصر j ام بردار پسماند $v-U\beta$ می باشد. برای حداقل کردن Φ نسبت به β قدم‌های زیر باید برداشته شود

قدم صفر) با مقدار دلخواه $\beta^{(k)}$ شروع می کنیم،

قدم اول) مقدار $\delta^{(k)}$ را چنان پیدا کنید که برای $\theta>0$ و به‌اندازه کافی کوچك $\Phi(\beta^{(k)}+\theta\delta^{(k)})\leq\Phi(\beta^{(k)})$ باشد.

قدم دوم) مقدار $\theta^{(k)}\geq 0$ را انتخاب می کنیم که بیشترین کاهش در Φ را سبب شود.

قدم سوم) قرار می دهیم $\beta^{(k+1)}=\beta^{(k)}+\theta^{(k)}\delta^{(k)}$.

زمانیکه $h=0$ ومجموع روی j درتابع $\Phi(\beta)$ (۱۹) برابرصفراست درنتیجه این ساده‌سازی قدم های فوق دقیقا " منطبق برالگوریتم (1978) Sinclair و Bartels, Conn می شود. سایرمسائل مرتبط راباتعدیل $\Phi(\beta)$ بایك مقدار μ برای حصول خانواده‌پارامتری توابع خطی تیکه‌ای $\Phi_\mu(\beta)$ میتوان حل نمودوبامراحل زیر جواب حداقل راپیداکرد.

قدم صفر) مقدار $\mu>0$ و $\beta=\beta(0)$ راانتخاب می کنیم.

قدم اول) $\Phi_\mu(\beta)$ رانسبت به β برمبنای روش فوق حداقل می کنیم.

قدم دوم) اگرشرط توقف لازم برآورده‌می شودتوقف می کنیم یادرغیراینصورت $\mu=\mu/10$ قرارداده‌وبه قدم اول بازمی گردیم.

نوآوری این مقاله درقراردادن گروه‌وسیعی ازمسائل درقالب دوالگوریتم فوق می باشد. این روش براحتی به الگوهای بانرم محدودنیزقابل بسط است.

(1980) Steiger و Bloomfield یك روش نزولی برای رگرسیون چندمتغیره نرم L_1 پیشنهادمی نمایند. الگوریتم آنان همچنین در (1983) Steiger و Bloomfield نیزتوضیح داده‌می شود. دربعضی ازمراحل این الگوریتم مرتبط باالگوریتم های (1940) Singleton و(1967b) Usow میباشد. مبنای این روش جستجوبرای یافتن مجموعه m مشاهده‌است که‌برروی رگرسیون نرم L_1 بهینه واقع می شوند، این مجموعه بابهبودپیاپی پیدامی شود، درهردوریك نقطه ازمجموعه فعلی بعنوان یك نقطه مناسب برای حذف مشخص می شود. این نقطه سپس با بهترین گزینه جایگزین می شود، نوآوری این روش در ارائه‌روش کارآ برای پیداکردن جایگزینی بهینه ویك روش‌رهگشابرای مشخص کردن نقطه‌ای است که‌بایدازلو لاحذف شود.

$x_1^T, \ldots, x_m^T$ رابعنوان ردیف های ماتریس طرح متغیرهای مستقل که منصوب به‌مجموعه فعلی نقاط $1, \ldots, m$ می باشندرانامگذاری نموده و x_m^T را برای جایگزینی درنظرمی گیریم. قرارمی دهیم:

$$y_i = x_i^T \beta \qquad i=1, \ldots, m-1 \tag{۲۰}$$

β رامجددا " تعریف می کنیم:

$$\beta = \beta^0 + t\delta \tag{۲۱}$$

که β^0 یك عضودلخواه ازمجموعه است وبردار δ دردستگاه زیرصدق می کند

$$x_i^T \delta = 0 \qquad i=1,\ldots,m-1 \tag{۲۲}$$

باداشتن این مجموعه نقاط ،مقداربهینه S راباحداقل کردن تابع زیرنسبت به t میتوان بدست آورد

$$\sum_{i=1}^{n} |y_i - x_i^T(\beta^0 + t\delta)| \tag{۲۳}$$

باجابجائی جملات داریم ،

$$\sum_{i=1}^{n} |w_i||r_i - t| \tag{۲۴}$$

که $w_i = x_i^T\delta$ و $r_i = (y_i - x_i^T\beta^0)/(x_i^T\delta)$ مقدار t راتوسط روش میانه وزنی لاپلاس میتوان محاسبه نمود. Bloomfield , Steiger یك تعدیل وزنی روش مرتب کردن سریع جزئی (1971) Chamber رابطورکارا برای پیداکردن مقدار t پیشنهادمی نمایند. نقطه داده آماری مرتبط بامیانه وزنی جایگزین x_k^T می شود. بااستفاده از y_i از(۲۰) ومعادله اضافی m ام،مقدار β^0 محاسبه می شود. بردار δ بامضربی از(۲۲) تعیین می شود. مجموعه‌جدیدمقادیر پارامترهابااستفاده‌از (۲۱) محاسبه می شوندحال یك نقطه بایدحذف شود. Steiger و Bloomfield یك راه‌مطمئن برای شناسائی این نقطه ارائه نمی نمایند. آنهایك روش رهگشابرمبنای گرادیان توصیه‌نموده ومقدارزیررابکارمی گیرند :

$$\rho = \frac{\left| \sum_{i:r_i<0} w_i - \sum_{i:r_i>0} w_i \right| - \sum_{i:r_i=0} w_i}{\sum_{i=1}^{n} w_i} \tag{۲۵}$$

بکبار ρ برای هرنقطه کەنامزدحذف شدن می باشد محاسبه‌گردیده‌ونقطه‌ای حذف می شود که ρ آن بیشترین مقداررادارد.

برای شروع الگوریتم هرمجموعه‌ای از m ردیف X رابا β^0 مناسب می توان انتخاب نمود. متغیرهارامرحله‌به مرحله‌اضافه‌نموده تایك برازش‌برای β^0 ومجموعه m نقطه‌مرتبط باآن بدست آید. درهرمرحله‌واسطه‌ای ،برازش شامل k متغیرکه $0 \le k < m$ ومجموعه k نقطه باپسماندصفرمربوطه‌می باشد. باواردکردن یك متغیرجدیدبرازشـــ را بهبودمی دهیم وبدین ترتیب k به k+1 افزایش می یابد. درهرمرحله ،مقدار ρ

است که تعیین می کندکه‌باالگوی بزرگتربرویم یا الگوی فعلی رابهبودبخشیم ،فرض کنیددر مرحله فعلی با k متغیرسروکارداریم ۰اگرمقدار ρ برای متغیر P ازهمه‌بیشتر است و P متعلق‌به مجموعه k متغیربا لانیست ، k به $k+1$ افزایش یافته ومتغیر p واردمی شودوبرازش رابا $k+1$ متغیربهبودمی بخشیم ،اگر p در مجموعه فعلی k متغیرمی باشدو درحداکثرمقدارخودمی باشد ،نقطه‌مرتبط به P حذف شده وهمانطورکه گفته شدجایگزین می شود ۰دراین مقاله رابطه این الگوریتم بابرنامه‌ریزی خطی نیزموردبحث قرارمی گیرد ۰

(1984) Steiger و Seneta یك الگوریتم برای جواب نرم L_1 دستگاه معادلات اندکی فوق تعیین پیشنهادمی نمایند ۰روش آنان برمبنای الگوریتم (1980) Steiger و Bloomfield که‌ذکرگردیدمی باشد ۰اگر m نزدیك به n باشداین الگوریتم سریعترازالگوریتم فوق است ۰اگرداشته باشیم

$$(x_i, y_i) \in R^{m+1}, \quad i=1,\ldots,n \quad k=n-m, \quad X=(x_1,\ldots,x_n)^T$$

الگوریتم آنان رامیتوان ازقرارزیرخلاصه نمود:

قدم‌اول) ردیف های $(X \mid y)$ راشماره‌گذاری مجددنموده بطوریکه X_N ، m ردیف ته X معکوس پذیرباشد ۰دستگاه k معادله‌ای خطی $X_N{}^T N = -X_T{}^T$ رابرای N حل می کنیم که X_T بیانگر k ردیف بالای X می باشد ۰

قدم‌دوم) بگذارید $D=(A \mid c)$ که $A=(I \mid N)$ باانداره $k \times n$ و $c=Ay$ ۰

قدم سوم) بگذارید $\sigma=(1,\ldots,k), \quad \sigma^c=(k+1,\ldots,n)$ ۰

قدم چهارم) جایگزاری کنید $r_{\sigma(i)}=b_i \quad i=1,\ldots,k$

$r_{\sigma^c(i)}=0 \quad i=k+1,\ldots,n$

قدم پنجم) بگذارید $I=\{i \mid 1 \le i \le k, c_i=0\}$

قدم ششم) حلقه تکراربرای $j=1$ m

بگذارید $v_i=D_{i\sigma^c(j)}$ برای $j=1,\ldots,k$ بگذارید

$$M=\{i \mid \operatorname{sgn}(c_i) \neq \operatorname{sgn}(v_i)\} \qquad J=\{1,\ldots,k\} \setminus N \setminus I$$

بگذارید

$$\beta_j = \frac{\left|\Sigma_M |v_i| - \Sigma_J |v_i|\right| - 1 - \Sigma_I |v_i|}{1+\sum_{i=1}^{k} |v_i|}$$

پایان حلقه

قدم هفتم) بگذارید $S=\{1,...,m\}$

قدم هشتم) β_q رابعنوان $\max_{s}\{\beta_j\}$ انتخاب کنید.

قدم نهم) اگر $\beta_q > 0$ است به قدم 11 می رویم

قدم دهم) اگر $\prod_{i=1}^{k} r_{\sigma(i)}=0$ ،مسئله دیژنره است وتوقف کنید.

درغیراینصورت $y_{\sigma c(i)}=\sum_{j=1}^{m} x_{\sigma c(i)j}\theta_j$ رابرای θ حل کنیدوتوقف کنید.

قدم یازدهم) بگذارید $v_i=D_{i,\sigma c(q)}$ برای $i=1,...,k$.

قدم دوازدهم) مقدار $t\hat{}$ برابربامیانه وزنی $c_1/v_1,...,c_k/v_k,0$

با وزن‌های $|v_1|,...,|v_k|,1$ راحساب کنید.

قدم سیزدهم) اگر $t\hat{}=c_p/v_p \neq 0$ به قدم شانزدهم می رویم.

قدم چهاردهم) بگذارید $S=S\setminus\{q\}$ اگر $S=\emptyset$ بەقدم دهم می رویم.

قدم پانزدهم) به قدم هشتم می رویم.

قدم شانزدهم) ردیف p ام D رابر $D_{p\sigma c(q)}$ تقسیم می کنیم.

قدم هفدهم) برای $i\neq p$ در D بگذارید ردیف i = ردیف i - ردیف p * $D_{i\sigma c(q)}$

قدم هجدهم) $\sigma(p)$ و $\sigma c(q)$ راجابجامی کنیم

قدم نوزدهم) بگذارید $r_{\sigma(i)}=b_i$ برای $i=1,...,k$ وقراردهید

$r_{\sigma c(q)}=0$.

قدم بیستم) به قدم پنجم بروید.

Seneta (1983) استفاده مکررمیانه وزنی رابرای برآورد بردارپارامترهادر الگوی خطی کلاسیك وقتی که معیار برازش نرم L_1 وهمچنین معیارکوشی می باشدرامرور می نمایند.

Wesolowsky (1981) یك الگوریتم برای رگرسیون چندمتغیرەنرم L_1 برمبنای نظریه نزول روی لبەچندوجهی تابع ابژکتیوارائه می نماید. این الگوریتم رابطه نزدیکی باالگوریتم های (1930) Rhodes و(1978) Sinclair و conn

Bartels که قبلا" توضیح داده شدنددارد. رگرسیون چندمتغیره راهمانندقبل در نظربگیرید. مجموعەای از m نقطه $(x_{j1}{}^{I},...,x_{jm}{}^{I},y_j{}^{I})$ راانتخاب کنید دستگاه معادلات زیرامیتوان برای یك مجموعه جواب یکتابرای ضرائب محاسبه نمود.

$$y_j{}^I - \sum_{h=1}^{m} \beta_h x_{jh}{}^I = 0, \quad j=1,\dots,m \qquad (۲۶)$$

یك لبه ازهرزیرمجموعه J با $m-1$ معادله شکل می پذیرد ، برای حداقل کردن در امتدادیك لبه فرض می کنیم

$$Y_I = \begin{bmatrix} y_1{}^I \\ \vdots \\ y_m{}^I \end{bmatrix}, \quad X_I = \begin{bmatrix} x_{11}{}^I & \dots & x_{1m}{}^I \\ \vdots & & \vdots \\ x_{m1}{}^I & \dots & x_{mm}{}^I \end{bmatrix},$$

$$Y_J = \begin{bmatrix} y_1{}^J \\ \vdots \\ y_{m-1}{}^J \end{bmatrix}, \quad X_J = \begin{bmatrix} x_{11}{}^J & \dots & x_{1m}{}^J \\ \vdots & & \vdots \\ x_{m-1,1}{}^J & \dots & x_{m-1,m}{}^J \end{bmatrix} \qquad (۲۷)$$

اگر β_p باحذف β_p از β بدست می آیدو x_p ستون P ام X_J و X_{pJ} باحذف x_p از X_J ایجادشود ، پس برای یك β_p داده شده‌می توان نشان دادکه

$$\beta_p = X_{pJ}{}^{-1} Y_J - \beta_p X_{pJ}{}^{-1} x_p \qquad (۲۸)$$

فرض کنیدعناصر β_p برابرباشندبا :

$$\beta_q = r_q - s_q \beta_p \quad \text{for } q \neq m \qquad (۲۹)$$

جایگزینی درتابع ابژکتیونرم L_1 می دهد

$$\min_{\beta_1} : \sum_{\substack{i=1 \\ \sum_{q\neq p}^{m} s_q x_{iq} \neq x_{ip}}}^{n} \left| x_{ip} - \sum_{q \neq p}^{m} s_q x_{iq} \right| \left| \frac{y_i - \sum_{q\neq p}^{m} r_q x_{iq}}{x_{ip} - \sum_{q=2}^{m} s_q x_{iq}} - \beta_1 \right|$$

حال قدم های زیررابایدبرداشت

قدم اول) بگذارید $k=1$ و $l=0$ ، مقادیراولیه‌ای برای $\beta_1, \dots, \beta_m$ انتخاب می کنیم. مقادیرحداقل مربعات یك امکان می باشد ، فرض کنید $I(1)=\{j_1(1), \dots, j_m(1)\}$ یك مجموعه m نقطه‌ازداده های آماری باشدکه به‌ترتیب زیرانتخاب شده‌اند ، نقطه‌ای که دارای کوچکترین پسماندمربع می باشدراهرباربه‌شرط اینکه $X_I(1)$ غیرمفردباشدانتخاب می کنیم ، مقدار $\beta_q(1)$, $q = 1, \dots, m$ راباحل $Y_I(1)=X_I(1)\beta$ پیدامی کنیم. جایگذاری می کنیم

$$I(k)=(j_1(k), \dots, j_m(k)); \quad J=(j_2(k), \dots, j_m(k))$$

قدم دوم) بگذارید $k=k+1$ مقدار β_p برای کوچکترین p راباا ستفاده از روش میانه وزنی محاسبه می کنیم ۰ $\beta_p(k)=\beta_p$ قرار می دهیم وفرض کنید i اندیسی باشدکه میانه وزنی پائینی β_p رابرای کمترین مقدار ممکن p مشخص نماید ۰

قدم سوم) الف)اگر $0 = \beta_p(k)-\beta_p(k-1)$ واگر $1>m$ باشدبه قدم چهارم میرویم درغیراینصورت قرارمی دهیم $I=(j_2(k-1),\ldots,j_m(k-1),i)$ و $l=l+1$ $\beta_q(k)=\beta_q(k-1)$ برای تمام q هاوبه قدم دوم می رویم۰

ب) اگر $0\neq\beta_p(k)-\beta_p(k-1)$ باشدقرارمی دهیم $l=0$ ،مقدار $\beta_q(k)$ رابرای $q\neq p$ ازرابطه (۲۹) محاسبه می کنیم ۰قراردهید $I(k)=(j_2(k-1),\ldots,j_m(k-1),i)$ وبه قدم دوم بروید ۰

قدم چهارم)مقدار $\beta_q(k)$ برای $q\neq p$ راازرابطه (۲۹) حساب نموده و $\beta^*=\beta(k)$ قرار داده‌وتوقف می کنیم۰

دراین مقاله Wesolowsky همچنین مسئله همخطی راموردبحث قرارداده‌وراه‌حل مناسب راارائه می نماید ۰

(1983) Sposito و Josvanger الگوریتم Wesolowsky رابرای الگوی رگرسیون خطی درپارامتری ساده تعدیل می نمایند ۰این تعدیل روش دیگری برای مرتب کردن مشاهدات بجای پشت سرهم قراردادن ترتیبی همه آنهابرای محاسبه‌مقدار میانه وزنی لازم می باشد ،آنها مقادیرکوچکترکه‌بایدمرتب شوندرابا وزنهای مربوطه در پائین $\beta_1(k-1)$ قرارداده ومقادیربزرگتریامساوی رابا لای آن قرارمی دهندوسپس نامساوی (۴) میانه وزنی راآزمایش می کنند ۰اگرنامساوی هابرقرارنشدند تغییرات لازم اعمال می گردد ۰مشخصا " ،اگرقسمت راست بیش ازحدوزن یافته است ،وزن مرتبط با کوچکترین عامل مرتب شونده‌به سمت چپ منتقل می شودوآزمایش مجددا " تکرارمی شود ۰یك برنامه کامپیوتربرای این الگوریتم همچنین توسط آنهاارائه گردیده‌است ۰

روش گرادیان تعمیم یافته Clarke (نگاه کنیدبه، (1983) Clarke) یك روش عمومی برای بهینه سازی مسائل وتوابع ناهموارمی باشد(نگاه کنیدبه (1986) Womersley و Pruess و Osborne)،یك زیرطبقه این روش بهنام گرادیان ساده شده که در(1985) Osborne توضیح داده شده‌است یك الگوریتم

عمومی است که الگوریتم های برنامه‌ریزی خطی ،بهینه سازی توابع محدب چندوجهی و بهینه سازی خطی تیکه ای را شامل‌می شود٠روش گرادیان ساده‌شده‌حالت خاصی از روش نزولی است که دارای دوخصوصیت عمده میباشد،مشخص کردن جهت وبرداشتن قدمی درجهت مزبوربرای کاهش مقدارتابع (همچنین نگاه کنیدبه (1975) Osborne و Anderson و (1985) Watson و Osborne و(1985,87) Osborne) الگوریتم های (1978) Sinclair و Conn و Bartels و(197a) Kung و Frome و Armstrong و(1980) Steiger و Bloomfield همگی حالات خاص ازروش گرادیان ساده شده می باشند٠

(1987) Yamamoto و Kato و Imai یك الگوریتم زمان خطی برای محاسبه کردن رگرسیون خطی دوپارامتری نرم L_1 بابکارگیری روش هرس ارائه می دهند٠ازآنجائیکه جواب بهینه درصفحه $\beta_0 x \beta_1$ برروی تقاطع خطوط داده های آماری می افتد،بنابراین ،درهرمرحله یك مجموعه ازخطوط مزبورکه جواب بهینه‌راتعیین نمی کندحذف می شوند،دراین مقاله توضیح جبری مسئله نیزارائه می گردد٠

(1987) Pilibossian نیزیك الگوریتم مشابه با(1958) Karst برای رگرسیون نرم L_1 خطی دوپارامتری ساده پیشنهادمینماید٠(1987a,b , 88a,b) Bidabad روشهائی نزولی برای رگرسیون های نرم L_1 ساده وچندمتغیره‌پیشنهاد می نماید٠این الگوریتم هابابهبودهای زیادی درفصل بعدموردبحث قرارخواهندگرفت و یك الگوریتم جدیداصلاح شده پیشنهادخواهدشد٠

۲-۳- الگوریتم های نوع سیمپلکس

منشاء بکارگیری برنامه‌ریزی خطی درحل مسئله نرم L_1 رامیتوان درمقاله (1888) Edgeworth پیدانمود٠(1950) Harris ابرازمی داردکه‌مسئله برآوردنرم L_1 درارتباط بابرنامه‌ریزی خطی است (1955) Ferguson و Cooper و Charnes این مسئله رادریك الگوی برنامه‌ریزی خطی فرموله کردند٠این مقالهاولین کاری است که برنامه‌ریزی خطی رادراین موضوع استفاده‌می نماید٠برنامه‌ریزی خطی رامیتوان به شکل زیربرای مسئله برآوردنرم L_1 بکارگرفت :

$$\begin{aligned}&\min_{\beta}: \ 1_n^T(w+v)\\&\text{s.to: } X\beta+I_n(w-v)=y\\&\qquad w,v\geq 0\\&\qquad \beta \text{ unrestricted in sign}\end{aligned} \qquad (۳۱)$$

که 1_n برداریك باندازه $nx1$ و I_n ماتریس یکه $n \times n$ می باشد. بردارهای v و w بهاندازه $n \times 1$ می باشندکه میتوان آنهارابترتیب بعنوان انحرافات با لاوپایین ازفوق صفحه رگرسیون درنظرگرفت این مسئله دارای n قیدتساوی با $m + 2n$ متغیرمی باشد. زمانی کهمقدار n بزرگ است این فرمولاسیون عموماً" فضای حافظه وزمان زیادی برای محاسبه لازم دارد.

Wagner (1959) نشان می دهدکهفرموله کردن رگرسیون نرم L_1 را میتوان بهمسئله برنامهریزی خطی با m قیدتساوی کاهش داد. بدین ترتیب، این فرمولاسیون ثانویه n معادلهفرم اولیه رابه m معادله فرم ثانویه کاهش می دهدوبطور قابل ملاحظهای فضای حافظه وزمان محاسبه راکم می نماید.

Fisher (1961) فرمولاسیون برآوردنرم L_1 رادررابطه بافرم اولیه برنامهریزی خطی بررسی می نماید. Barrodale و Young (1966) یك الگوریتم سیمپلکس تعدیل شده برای تعیین بهترین برازشی بهمجموعهای ازدادههای گسسته بامعیارنرم L_1 رابسط می دهد، این روش بهزبان Algol ارائه می شود(برای بحث دراین روش، نگاه کنیدبه (Sposito (1975 و Mc Cormick ((1967).

Davies (1968) استفادهبرآوردرگرسیون نرم L_1 رانشان می دهد. Rabinowitz (1969) نیزکاربردبرنامهریزی خطی رادراین زمینه بحث می نماید.

Crocker کاربردمعیارنرم L_1 رافقط برای اجتناب ازبرآوردضرائب منفی ناخواسته که دررگرسیون حداقل مربعات اتفاق می افتدتوصیه می نماید. همخطی یکی ازاین حالات است که منتج بهاین نتیجهمی شود، (Ben ـ Israel (1969 و Robers با استفاده ازبرنامهریزی خطی فاصلهای یك الگوریتم برای حل مسئله برآوردنرم L_1 پیشنهاد می‌نمایند. (Rabinowitz (1970 و (Weil (1970 و Shanno ارتباط برنامهریزی خطی ومسئله تقریب راموردبحث قرارمیدهند. (Barrodale (1970

برازش نرم L_1 خطی و غیرخطی را درمورد داده‌های آماری گسسته وپیوسته خلاصه می نماید.

(1973) Young و Kiountouzis و Spyropoulos دوالگوریتم برای برازش توابع کلی ومخصوصاً " الگوریتم سریع با حداقل نیاز حافظه برای برازش چند جمله‌ای برمبنای خواص جبری، فرمولاسیون برنامه‌ریزی خطی پیشنهادمی نمایند . (1973)

Robers و Robers نوع خاصی ازروش عمومی (1969) Ben - Israel و Robers که‌مشخصاً " برای مسئله نرم L_1 طراحی‌شده‌است راارائه می نمایند.برنامه فورترن مربوطه‌نیزهمراه است .

(1973) Roberts و Barrodale تعدیلی ازروش سیمپلکس که‌احتیاج به‌میزان حافظه کمتری داردوباحذف کردن تارك های سیمپلکس خیلی کاراترازروش سیمپلکس معمولی می باشدراپیشنهادمی نمایند.اگربردار β رابه‌عنوان اختلاف دوبردارغیرمنفی c و d تعریف کنیم ،فرمولاسیون آنان رامیتوان به‌شکل زیرنوشت :

$$\begin{aligned} &\min_{c,d}: \ 1_n^T(w+v) \\ &\text{s.to: } X(c-d)+I_n(w-v)=y \\ &\qquad w,v,c,d\geq 0 \end{aligned} \tag{۳۲}$$

بعلت رابطه میان متغیرها ،میتوان بااستفاده‌ازفضای بردارهافقط به‌میزان (n+2)x(m+2) شامل برچسب ها برای بردارهای پایه وغیرپایه محاسبات را انجام داد.اگرتمام y_i هاغیرمنفی باشندیك پایه اولیه را w میتوان فرض نمود.اگر یك y_i منفی است ،علامت ردیف مربوطه‌راتغییرمی دهیم وستون یك عضومربوطه به v رابعنوان قسمتی ازپایه درنظرمی گیریم .این الگوریتم دردومرحله‌شکل میگیرد.مرحله اول انتخاب ستون لولادرحین m دورتکراراول برای عناصربرداری، c_j و d_j برمبنای حداکثرهزینه‌های نهائی غیرمنفی مربوطه محدودمی سازد.برداری که پایه راترك می کند سبب حداکثرکاهش درتابع ابژکتیومی شود.بنابراین عنصرلولالزوماً " همانند سیمپلکس معمولی نیست .مرحله دوم w_i یا v_i غیرپایه رابا w_i یا v_i پایه‌جابجا می نماید.بردارهای پایه درارتباط با c_j و d_j مجازبه ترك پایه نمی باشند. الگوریتم زمانی متوقف می شودکه‌تمام هزینه‌های نهائی غیرمثبت باشند(نگاه‌کنیدبه (1980) Gentle و Kennedy) .یك برنامه فورترن برای این روش توسط

Willms (1983) نوشته می شود(Barrodale و Roberts (1974)
و Peters الگوریتم هائی همراه با برنامه‌های کامپیوتر برای به‌نگام کردن جواب مسئله وقتی که یك ستون یا ردیف از X ویا y حذف ویا اضافه می شود ارائه می نمایند. این الگوریتم ها همگی بر اساس روش (1973, 74) Roberts و Barrodale می باشند.

Abdelmalek (1974) یك الگوریتم سیمپلکس ثانویه را برای مسئله نرم L_1 بدون استفاده از متغیرهای مصنوعی شرح میدهد. در مورد این الگوریتم احتیاج نیست که شرط Haar (نگاه کنید به (1985) Osborne و (1961) Moroney) برقرار باشد. این الگوریتم در زمان انتشار خیلی کارا بنظر می آمد. یك الگوریتم سیمپلکس ثانویه بهبود یافته برای تقریب نرم L_1 توسط (1975a) Abdelmalek پیشنهاد میشود، در این الگوریتم، دورهای تکرار میانی معینی حذف می شوند و در حالت مسائل ill -- conditioned ماتریس پایه را می توان از طریق فاکتورگیری به یك ماتریس مثلثی تبدیل کرد و جواب با ثبات را تامین نمود. (1980a) Abdelmalek الگوریتم قبلی خود را با تجزیه مثلثی بهبود می بخشد. یك ترجمه از این الگوریتم به زبان فورترن در (1980a) Abdelmalek ارائه می شود. (1978) Kennedy و McCormick و Sposito اکثر آثار دیگران در مورد برآورد نرم L_1 شامل طرح مسئله، فرمولاسیون برنامه ریزی خطی، الگوریتم های محاسباتی کارا و خواص برآوردکننده‌ها را خلاصه می نمایند.

(1978) Kung و Armstrong یك الگوریتم برای رگرسیون نرم L_1 دو پارامتری ساده پیشنهاد می نمایند. این روش دارای مشخصه برنامه ریزی خطی الگوریتم (1973) Roberts و Barrodale می باشد. یك برنامه فورترن نیز ارائه شده است.

(1979a) Kung و Frome و Armstrong روش تجزیه (مثلثی پائینی - بالائی) LU متعلق به (1969) Golub و Bartels را در ترمیم پایه فعلی روش سیمپلکس تجدید نظر شده بکار می‌بندند. یك ترجمه فورترن نیز همراه می باشد. (1979) Godfrey و Armstrong نشان می دهند که روش اولیه (1973) Roberts و Barrodale و روش ثانویه (1975)

Abdelmalek اصولا "معادل می باشند ،بایك پایه اولیه داده شده برای دو روش ،آنها نشان می دهندکه هردوالگوریتم درهردوپایه های متناظرایجادمی نمایند ،تنها اختلاف درانتخاب پایه اولیه وقواعدرهگشابرای رسیدن به جواب می باشد. (b 1982)

Kung و Armstrong یك فرمولاسیون ثانویه برنامه ریزی خطی برای مسئله ارائه می نمایند ،ورودهای مختلف، وروشهای معین کردن پایه اولیه موردملاحظه قرارمی گیرند. نشان داده شده است که اگریك جواب شدنی ثانویه خوب وجودداشته باشدروش ثانویه برتراز اولیه می باشد(نگاه کنیدهمچنین به (1980) Steiger). (1980)

Taylor و Banks تعدیلی ازالگوریتم (1973) Roberts و Barrodale راپیشنهاد می نمایند ،تابع ابژکتیوبرای دربرگرفتن مقدارعناصربردارهای جواب وخطا تغییرمی یابد ،برای بحث کلی درموردسیمپلکس برای برنامه ریزی خطی تیکه ای رجوع کنید به (1985a , b) Fourer وبرای مروری به این مسئله درارتباط بانرم L_1 نگاه کنیدبه (1986) Fourer (1987). Wellington و Narule یك الگوریتم برنامه ریزی خطی کارا برای جلدوررگرسیون چندمتغیره خطی نرم L_1 و L_∞ پیشنهادمی نمایند ،این الگوریتم ازساختارمخصوص وشباهت های بین این دومسئله بهره می گیرد.

(1987) Seiford و Brennan یك تعبیرهندسی برنامه ریزی خطی رادررگرسیون نرم L_1 بسط میدهند ،آنهایك دیدهندسی ازپروسه جواب رادرفضای مشاهدات ارائه می نمایند. (1987) McConnel نشان می دهدکه چگونه روش ژاکوبین های صفرکه برای بهینه سازی مسائل برنامه ریزی کوا راتیك مورداستفاده واقع می شود رامیتوان برای حل مسئله خاص برنامه ریزی خطی مرتبط با محاسبه تقریب نرم L_1 گسسته خطی بکاربرد. برای امکان بکارگیری سایرانواع روشهای حل برنامه ریزی خطی مانندحل Karmarkar درمسائل نرم L_1 نگاه کنیدبه (1986) Meketon.

۳-۳- سایرالگوریتم ها

این گروه شامل الگوریتم هائی می شودکه دردوقسمت قبل طبقه بندی نشدند.

(1964c) Rice روش دونیم کردن رابرای رگرسیون نرم L_1 بکارمی برد.

دراین روش درهرقدم دامنه S به دوقسمت تقسیم می شودوقسمت مناسب برای دوربعد انتخاب می گردد. وقتی آخرین قسمت کمتر ازمقدارکوچك ازقبل تعیین شده باشدجواب حاصل گردیده است (برای بحث برروی روش دونیم کردن رجوع کنیدبه (1989) Bidabad)

(1971) Abdelmalek الگوریتمی برای برازش توابع به مجموعه نقاط گسسته وحل دستگاه فوق تعیین معادلات خطی راتوسعه می دهد. این روش براساس تعیین جواب نرم L_1 بعنوان حالت حدی تقریب نرم L_p وقتی p ازسمت راست درحدبه یك میل کندمی باشد. بنابراین ، این روش با حل یك دنباله ازمسائل غیرخطی جواب یك مسئله خطی رامی یابد.

(1973) Schlossmacher برآوردهای پارامتری رگرسیون نرم L_1 رابایك روش حداقل مربعات وزنی تکراری بدست می آورد. به جای حداقل کردن مجموع قدرمطلق خطاها ، وی مجموع وزنی مربع خطاهارابا وزن $1/|u_i|$ حداقل می کند. یکباررحداقل مربعات درموردمسئله بکارگرفته شده وپسماندهامحاسبه می گردند. قدرمطلق معکوس خطاهامجددا" به عنوان وزن های مربوطه برای حداقل کردن مجموع خطاهای وزنی دردوربعد استفاده می شوند(همچنین نگاه کنیدبه (1977) Welsh و Holland).

(1974) Fair ملاحظه نمودکه مقادیربرآوردشده β بعدازدوردوم یاسوم تغییر نمی نمایند. درحالاتی که یك پسماندصفراست ، ادامه روش ناممکن می گردد. این مسئله همچنین توسط (1977) Kennedy , Gentle , Sposito و(1988) Rouhi Christensen , Soliman موردبحث قرارمی گیرد. همگرائی مطلق این الگوریتم اثبات نشده ولی یك آزمایشی غیرهمگراتابحال گزارش نگردیده است.

(1988) Rouhi و Christensen و Soliman شبه معکوس چپ (نگاه کنیدبه (1978) Dhrymes برای توضیح این معکوس) رابرای حل رگرسیون نرم L_1 خطی عمومی بکارمیگیرند. براساس این روش ، جواب حداقل مربعات رابا استفاده ازشبه معکوس چپ یاتقریب حداقل مربعات $\hat{\beta}=(X^TX)^{-1}X^Ty$ بدست می آورند. پسماندهارابه شکل $u=|y-X\hat{\beta}|$ که u برداری $n\times 1$ است حساب می کنیم. m مشاهده باکوچکترین مقادیرقدرمطلق پسماندهاراانتخاب کرده وماتریس ها راطوری افرازمی کنیم که مشاهدات منتخب دربالاقرارگیرند.

$$u=\begin{bmatrix}u^{\wedge}\\ --\\ u^{\sim}\end{bmatrix}, \quad X=\begin{bmatrix}X^{\wedge}\\ --\\ X^{\sim}\end{bmatrix}, \quad y=\begin{bmatrix}y^{\wedge}\\ --\\ y^{\sim}\end{bmatrix}$$

برای پارش بالا $y^{\wedge}=X^{\wedge}\beta^{\wedge}$ رابه‌شکل $\beta^{\wedge}=X^{\wedge -1}y$ حل می کنیم • اگر چه‌این روش عملا" ساده‌است ،جواب آن همانندسایرروشهای دقیق نیست واثباتی ارائه‌نشده است که آیاجواب درهمسایگی جواب دقیق حل مسئله حداقل سازی نرم L_1 می باشدیا خیر •

کاربرد median polish (نگاه کنیدبه (1977) Tukey) و median polish -$\in$ درنرم L_1 توسط (1983) Bloomfield, Steiger و(1984) Kemperman و(1987a) Sposito (1987a,b) Bradu بحث می شود •

کاربردالگوریتم Karmarkar برای برنامه‌ریزی خطی ورابطه آن بانرم L_1 در(1985) Kim و Skarpness و Sherali آمده‌است •برای استفاده روش هموتوپی درنرم L_1 رجوع کنیدبه(1983) Garcia , Gould و(1987) Schellhorn یك‌الگوریتم برای تقریب نرم L_1 خطی برای توابع پیوسته را(1981) Watson ارائه می نماید(همچنین نگاه کنیدبه (1981) Baboolal , Watson)

۳-۴- مسئله مقداراولیه

چگونگی شروع الگوریتم هاتوسط نویسندگان زیادی موردبحث واقع شده‌است • انتخاب مقدار اولیه‌عامل مهمی درزمان اجرای الگوریتم های مختلف می باشد•به عبارت دیگر ، یك نقطه شروع خوب جواب راسریعتربدست می دهدوتعدادتکرارهاراکم می نماید•مقالات متعددی وجوددارندکه‌این مسئله رابرای الگوریتم های حداقل سازی نرم L_1 ملاحظه‌می‌کنند (1968) Duris , Sreedharan بطورمختصربه‌این مسئله اشاره می نمایند• (1976) McCormick , Sposito برآوردکننده‌حداقل مربعات رابرای ساختن یك نقطه شروع خوب الگوریتم (1973) Barrodale , Roberts استفاده نمودند•این مقداراولیه تعدادتکرارهارادرحداکثرحالات کاهش داد(1977) McCormick Sposito , Hand نشان می دهندکه کل زمان CPU لازم برای حصول ضرائب

رگرسیون بهینه درنرم L_1 رابطورکلی اگرنزدیکترین برآوردکننده به نرم L_1 مانند حداقل مربعات اول محاسبه شودسپس روش تعدیل شده(1973) Roberts و Barrodale رابکاربریم می توان کاهش داد ۰بحثی مشابه درموردبرآوردنرم L_∞ توسط (1980) Sposito , Hand ارائه می شود(1982) Armstrong و Sklar ابرازمی دارندکه بکارگیری پسماندهای حداقل مربعات برای ایجادیك شروع خوب برای الگوریتم (1987) Kung , Frome , Armstrong منتج به کاهش بسیارزیاددرحجم محاسبات می گردد ۰

۳ـ۵ـ برنامه های کامپیوتری ونرم افزارها

اگرچه بیشترنویسندگان برنامه های کامپیوتری برای الگوریتم های خودشان راکد کرده‌اندکه قبلا" به آنها اشاره شد ،برنامه‌های کامپیوتری دیگری نیزوجوددارندکه‌مسئله رگرسیون نرم L_1 راحل نموده وآمارهای لازم راحساب می کنند ،بعضی ازاین برنامه های ازپیش نوشته شده عبارتنداز IMSL (نگاه کنیدبه (1985) Rice) ، BLINWDR (نگاه کنیدبه (1987) Dutter)، ROBETH (نگاه کنیدبه (1987) Marazzi ، (1985) Marazzi , Randriam-iharisoa)و ROBSYS (نگاه کنیدبه (1987) Hardle) ۰ XploRe

ازآنجائیکه این نرم افزارهاویژگی های خاص خودرادارا هستندمابه جزئیات آنهانخواهیم پرداخت ۰ علاقه مندان میتوانندبه منابع ذکرشده رجوع نمایند ۰

۳ـ۶ـ مقایسه الگوریتم ها

بطورکلی ،مقایسه الگوریتم هایك کارسرراست نیست ۰ همانطورکه (1977) Dutter خاطرنشان می سازد ،عواملی چون کیفیت برنامه‌های کامپیوترومحیط محاسبه بایدمدنظرقرارگیرند ۰درحالت الگوریتم های نرم L_1 سه عامل مشخص تعداد مشاهدات ،تعدادپارامترهاوشرایط داده‌ها مهمترمی باشند(1977a,b) Sposito و Gentle و Kennedy و(1980a,b) Shier و Hoffman روشهائی برای ایجادداده‌های تصادفی آزمایشی بابردارهای جواب نرم L_1 دانسته را توضیح می دهند ۰ (1977) Gilsinn et al یك متدولوژی کلی برای

مقایسه الگوریتم های نرم L_1 راموردبحث قرارمی دهند(1977) Gentle
وو Kennedy خطای گردکردن دررگرسیون نرم L_1 رابررسی نموده ودوروش بـــرای
برخوردبا اشتباهات محاسبه پیشنهادمی‌نمایند(همچنین نگاه کنیدبه (1980) Sameh
Larson ,).

نویسندگان زیادی الگوریتم های خودرابا الگوریتم های پیشنهادشده قبلی مقایســـه کرده‌اند. جدول ۱ خلاصه‌ای ازالگوریتم های پیشنهادشده توسط نویسندگان مختلف راارائـه می‌نماید. لازم به ذکراست که ازآنجائیکه توزیع خطاهای رگرسیون الگوریتم های ارائـــه شده درجدول 1 متفاوت می باشد، ازاین جدول نبایدبطور مطلق نتیجه‌گیری ومقایسه‌کرد.

جدول ۱ : خلاصه ویژگی های الگوریتم های موجود

ref.	compared with	m range	n range	time/ performance
BCS	BR	2-8	201	roughly equal speed
AFK	BR	5-20	100-1500	30%-50% AFK is faster
A	BR	1-11	15-203	nearly equal speed
BS	BR	2-6	100-1800	BS is faster for larger n
W	AFK,AK	2-25	100-1000	W is faster for larger n and smaller m
SS	BS	4-34	10-50	SS is faster for m near n
AK	S	2	50-500	AK is faster
JS	Ak	2	10-250	JS is faster

n ≡number of observations.
m ≡number of parameters.
BCS≡Bartels,Conn,Sinclair(1978).
BR ≡Barrodale,Roberts(1973,74).
AK ≡Armstrong,Kung(1978).
S ≡Sadovski(1974).
AFK≡Armstrong,Frome,Kung(1979).
A ≡Abdelmalek(1980a,b).
BS ≡Bloomfield,Steiger(1980).
W ≡Wesolowsky(1981).
JS ≡Josvanger,Sposito(1983).
SS ≡Seneta,Steiger(1984).

(1976 a) Frome و Armstrong حداقل مربعات وزنـــی
تکراری (1973) Schlossmacher راباالگوریتم (1973) Roberts
و Barrodale مقایسه می کنندکه نتیجه برتری الگوریتم دوم بود.

Steiger و (1980) الگوریتم های Anderson , Steiger (1973)

(1973) و Bartels , Conn, Sinclair (1978) Bloomfield

Barradale , Roberts رامقایسه می نمایند۰این بررسی نتیجه‌گیری

می‌نمایدکه‌با افزایش تعدادمشاهدات ،الگوریتم BR دریک طبقه پیچیدگی متفاوت از

BCS و BS قرارمی گیرد۰هرسه الگوریتم برحسب تعدادپارامترهاخطی بوده و BS از

پیچیدگی کمتری نسبت به BCS برخورداراست ۰ پیچیدگی BS و BCS بر

حسب n خطی می باشند۰یك تمایل خفیف درموردکلیه الگوریتم هامشاهده شدکه‌برای

m های زوج نسبتا " سخت تربوده اندتا m فرد BR و BS بیشترین مشکل راباتوزیع

نرمال خطاهاداشتندوحداقل سختی راوقتی پارامترچگالی پارتومربوطه مساوی ۱/۲ بوده

باتوزیع پارتوداشتند۰

جدول ۲ : نیازحافظه بردارهابرای الگوریتم های منتخب

program name	ref.	required array storage	stopping constants
L1	BR	3n+m(n+5)+4	BIG=1.0E+75 TOLER=10**(-D+2/3) D=number of decimal digits of accuracy
L1	A	6n+m(n+3m/2+15/2)	PREC=1.0E-6 ESP=1.0E-4
L1NORM	AFK	6n+m(n+m+5)	ACU=1.0E-6 BIG=1.0E+15
BLAD1	BS	4n+2m(n+2)	-------------------
LONESL	S	4n	PREC=1.0E-6 BIG=1.0E+19
SIMLP	AK	4n	ACU=1.0E-6 BIG=1.0E+19
DESL1	JS	5n	TOL=1.0E-6

نگاه کنیدبه جدول ۱ برای اختصارات.

منبع: Gentle , Narula , Sposito(1987).

Gentle, Narula , Sposito (1987) مقایسات نسبتا " کاملی رابرای بعضی ازالگوریتم های نرم L_1 انجام میدهند. آنها این مقایسه رابه کدهائی که بـــرای رگرسیون فرم بی قیدخطی نرم L_1 موجودبودمحدودساختندجدول ۲ حافظه بردارهـــای لازم وثابت های توقف الگوریتم های مزبوررانشان می دهد.

درمطالعه آنها مسئله شامل مقادیرتصادفی یکنواخت (0,1) برای X ومقادیـــر تصادفی نرمال (0,3) برای جمله خطامی باشد. مقدار y بعنوان جمع متغیرهـــای مستقل وجمله خطامحاسبه گردید. خلاصه نتایج درجداول ۳ و ۴ بترتیب برای رگرسیون هـــای ساده وچندمتغیره نشان داده شده است . مقادیردرون سلول ها میانگین زمان CPU ، ۱۰۰ آزمایش، مقادیرداخل پرانتزهاحداکثرزمان CPU مربوطه برای ۱۰۰ آزمایش می باشد.

جدول ۳ : زمـــان CPU برای الگوی ساده

n	AK	JS	A	AFK	BS
100	0.021	0.023	0.094	0.034	0.023
	(0.03)	(0.04)	(0.21)	(0.06)	(0.04)
500	0.193	0.302	1.434	0.287	0.145
	(0.38)	(0.61)	(3.13)	(0.49)	(0.26)
1000	0.544	0.971	4.775	0.784	0.422
	(1.36)	(2.16)	(10.60)	(1.76)	(1.19)
5000	1.262	2.837	211.23*	1.614	+
	(24.58)	(48.88)	(----)	(31.22)	+

نگاه کنیدبه جدول ۱ برای اختصارات.

* میانگین ۳ اجرا.

+ موفق به محاسبه جواب صحیح نشد.

منبع : Gentle, Narula,Sposito (1987).

(1988) Narula , Sposito , Gentle نیزالگوریتم های رگرسیون هـــای نرم L_1 بی قیدخطی رامقایسه می نمایند. این بررسی بطورکلی خلاصه‌ای از (1987) Gentle , Narula,Sposito می باشد. نتایج بدست آمده‌کاملا" مشابه می باشد.

جدول ۴ : زمان CPU برای الگوی چندمتغیره (m =5,15)

n	m	A	AFK	BS
100	5	0.331 (0.53)	0.149 (0.23)	0.114 (0.17)
100	15	1.976 (2.73)	1.313 (1.70)	0.933 (1.38)
500	5	3.686 (5.47)	1.120 (1.81)	0.829 (1.22)
500	15	17.876 (23.4)	7.808 (10.1)	7.294 (9.13)
1000	5	13.211 (18.3)	2.930 (4.38)	+
1000	15	49.866 (72.7)	17.901 (24.0)	+
5000	5	248.91* (----)	34.311 (51.8)	+
5000	15	687.31* (----)	140.321 (160.1)	+

نگاه‌کنیدبه‌جدول ۱ برای اختصارات.

* میانگین ۳ اجراء.

+ موفق به محاسبه جواب صحیح نشد.

منبع : (1987) Gentle, Narula ,Sposito.

آنهانتیجه گیری می نمایندکه برنامه BS درمسائل کوچك عملکردخیلی خوبـــی داردولی درمسائل بزرگتربعلت انباشت خطای گردکردن موفق به محاسبه جواب صحیـــــح نمی شود. افزایش دقت برنامه کدشده برای اجتناب ازخطای گردکردن زمان اجـــــرا ءرا بالامی برد، بنابراین روشن نیست که چه برسرکارائی نسبی الگوریتم BS بعــــــد از این تعدیل خواهدآمد.

برنامه Wesolowsky درمطالعات آنها قابل استفاده نبودوازاین بابت حذف گردید. بدلیل برتری AFK به BR و AK به S كه‌درمطالعات قبلی شناخته شده بود، الگوریتم های BR و S راواردمقایسات نکردند.

با ملاحظه کلیه جنبه‌ها، آنها نتیجه‌گیری می نمایندکه AFK بهترین بنظر می رسد.

۷ـ۳ روشهای محاسباتی فرم غیرخطی

مجددا " فرض کنیدکه y, X, u, β همانندقبل تعریف شده‌اند. دررگرسیون نرم L_1 غیرخطی ، مسئله برآوردبردار β درالگوی غیرخطی زیراست

$$y_i = f_i(x_i,\beta) + u_i \qquad i=1,...,n;\ n\geq m \tag{۳۳}$$

که f_i تابع‌پاسخ و x_i ردیف i ام X می باشد. پارامترهای رگرسیون نرم L_1 باحداقل کردن مجموع زیربدست می آید.

$$\min_{\beta}: \sum_{i=1}^{n} |y_i - f_i(x_i,\beta)| \tag{۳۴}$$

تابع (۳۴) رامیتوان به شکل یك مسئله برنامه‌ریزی غیرخطی زیرنوشت .

$$\begin{aligned}
&\min_{\beta}: \sum_{i=1}^{n} w_i \\
&\text{s.to: } y_i - f_i(x_i,\beta) - w_i \leq 0 \\
&\qquad -y_i + f_i(x_i,\beta) - w_i \leq 0 \\
&\qquad\qquad w_i \geq 0 \\
&\qquad i=1,...,n
\end{aligned} \tag{۳۵}$$

دردودهه پیش الگوریتم های متعددی برای حل مسئله رگرسیون غیرخطی نرم L_1 پیشنهادشده‌است. این روشهارا می‌توان درسه گروه اصلی زیر طبقه‌بندی نمود(نگاه کنیدبه (1987 b) Gonin , Money وبرای سایرطبقه بندیهانگاه کنیدبه (1986) Watson و(1980) Mclean , Watson)

اولین گروه شامل روشهائی می شودکه فقط ازمشتق مرتبه اول استفاده می نمایند. در این الگوریتم ها مسئله غیرخطی اصلی به دنباله‌ای ازمسائل نرم L_1 خطی ساده می شوندکه هرکدام ازآنها با استفاده ازروش استانداردبرنامه‌ریزی خطی بطورخیلی کارآ قابل حل است.

این روش ها ازنوع Gauss - Newton می باشند. الگوریتم های اصلی که دراین گروه قرارمی گیرندتوسط نویسندگانی چون (1971) Osborne , Watson و(1977a ,b) Anderson , Osborne (1980) Sharger , Hill و(1980) Mclean , Watson و(1980) Osborne و Jittorntrum (1980) و(1980, 84a) Osborne و Watson (1982) Bartels,Conn و(1985) Madsen و Hald ارائه شده‌اند.

گروه دوم شامل روشهائی است که با استفاده ازمشتق مرتبه دوم، مسئله اصلی رابه دنباله‌ای ازمسائل حداقل سازی بدون قیدتبدیل می‌نمایند. مشتق ناپذیری تابع ابژکتیوازاین پس مانعی نخواهدبود. این روش به نام روش تابع پنالتی دربرنامه‌ریزی غیرخطی شناخته می شود. مقالاتی دراین زمینه توسط (1979) Vidyasagar , Dutta و El - attar (1981, 84) Fletcher و(1982) Zang و Tishler و(1984) Conn و(1987) Conn , Gould ارائه گردیده است.

درآخرین گروه تابع ابژکتیوراخطی نموده ولی تقریب های کوادراتیک رابرای تاثیر گذاشتن برانحنابکارمی‌گیریم (نگاه کنیدبه (1981) Murray , Overton و(1982) Overton و(1982) Bartels , Conn).

سایرویژگی های مسئله نرم L_1 غیرخطی توسط (1984a , b) Rice و (1978) Osborne , Watson (1979) Charalambous و (1979) Glashoff , Schultz و(1981a,b) Hald و(1982) Wagner و(1982,87) Waston و(1984) Yuan و Powell بحث می شود.

۸ـ۳ محاسبات نرم L_p

الگوی رگرسیون خطی را همانند قبل درنظر بگیرید ، برآوردنرم L_p β را میتوان بـــــا حداقل کردن مجموع توان p ام قدرمطلق خطاها بدست آورد ، بدین شکل :

$$\min_{\beta}: \sum_{i=1}^{n} \left| y_i - \sum_{j=1}^{m} \beta_j x_{ij} \right|^p \qquad (۳۶)$$

مسئله فوق را میتوان به شکل یك مسئله برنامه‌ریزی ریاضی تبدیل نمود ، بردار خطـــا را بعنوان تفاوت دوبردار غیرمنفی w و v که بترتیب بیانگر انحرافات مثبت ومنفـــی می باشندمی نویسیم ، به عبارت دیگر $u=w-v;\ w,v\geq 0$ مسئله تقریب نرم L_p به شکل زیرساده می شود (نگاه کنیدبـــــــه (Kiountouzis 1972)) ،

$$\begin{aligned} \min_{\beta}: & \sum_{i=1}^{n} (w_i^p + v_i^p) \\ \text{s.to:}\ & w_i - v_i + \sum_{j=1}^{m} \beta_j x_{ij} = y_i \\ & w_i, v_i \geq 0 \\ & \beta_j \text{ unrestricted in sign} \\ & i=1,\dots,n;\ j=1,\dots,m \end{aligned} \qquad (۳۷)$$

بایدیادآوری نمودکه‌این فرمولاسیون خیلی انعطاف پذیراست بطوریکه میتوان هرقیـــد دیگری رااضافه نمود(نگاه کنیدبه (Hart 1978) و Graves - Affleck , Money) مشخصه جالب دیگراین است که می توان الگورابه راحتی باحذف جمله مجموع در n قیداول وگذاشتن $f_i(x_i,\beta)$ به‌جای آن به فرم غیرخطی تعمیم دادبدین شکل ،

$$\begin{aligned} \min_{\beta}: & \sum_{i=1}^{n} (w_i^p + v_i^p) \\ \text{s.to:}\ & w_i - v_i + f_i(x_i,\beta) = y_i \\ & w_i, v_i \geq 0 \\ & \beta_j \text{ unrestricted in sign} \\ & i=1,\dots,n;\ j=1,\dots,m \end{aligned} \qquad (۳۸)$$

Valentine و Van Dine (1963), Aoki (1965) , Osborne , Watson (1967), Bartels , Golub (1968 a, b),Gustafson , Kortanek , Rom (1970), Barrodale , powell , Roberts (1972), Cline (1972,76) ,Duris , Temple (1973),Watson (1973),Barrodale ,phillips (1974,75) ,Boggs (1974), Fletcher , Grant , Hebden (1974a), Madsen (1975) , Abdelmalek (1975b,76,77a,b) , Conn (1975), Coleman (1978) ,Charalambous , Conn(1978) Bartels , Conn , Charalambous (1978),Armstrong , Kung (1979),Kling - man, Mote (1982),Bartels , Conn , Li (1987),Brannin- gan, Gustafson (1987).

۴ـ دستگاه معادلات همزمان

برآوردنرم L_1 به صورت گسترده برای الگوی رگرسیون تک معادله‌ای مطالعه‌گردیده وخواص آن بخوبی شناخته شده‌است ٠ولی علیرغم کاربرداقتصادسنجی زیادبرآوردنرم L_1 دردستگاه معادلات همزمان ،محققین معدودی دراین زمینه کارکرده‌اندکه آثارآنها در این قسمت خلاصه خواهدشد ٠معادله زیررابعنوان معادله اول دستگاه ساختاری درنظر بگیرید،

$$y = Y\theta + X_1\beta + u = [Y \mid X_1]\begin{bmatrix}\theta \\ --- \\ \beta\end{bmatrix} + u \equiv Z\alpha + u \qquad (۳۹)$$

که y بردارمتغیرهای درون‌زای تابع ، Y ماتریس متغیرهای درون‌زای مستقل، X_1 ماتریس متغیرهای برون‌زا ، θ و β بردارهای پارامترهای رگرسیون و u بردارخطاهای تصادفی می باشندفرم خلاصه شده Y به شکل زیرمی باشد،

$$Y = X\pi + v \qquad (۴٠)$$

حداقل قدرمطلق انحرافات مستقیم وغیرمستقیم (IDLAD , DLAD) مشابه‌با

نتیجه فرمولاسیون مسئله برآوردنرم L_p غیرخطی می باشد.

درحالت کلی رگرسیون نرم L_p روشهای محاسباتی بسیاری برای الگوهای خطـــی و همینطورغیرخطی وجوددارد(درموردجزئیات آن ،علاقه مندان میتوانندرجوع کنندبــه :

Descloux (1963), Rice (1964,69),Barrodale ,Young (1966),Sreedharan (1969,71),Ekblom , Henriksson (1969),Karlovitz(1970a,b),Barrodale , Roberts (1970), Barrodale ,Roberts, Hunt (1970),Fletcher, Grant ,Hebden (1971,74b), Kiountouzis (1972), Forsythe (1972) , Kahng (1972) ,Ekblom (1973a,b) , Anton,Duris (1973) , Watson(1973,77,78,84b,85a), Shisha (1974) , Merle , Spath (1974), Oettli (1975), Rey (1975), Mond , Schechter (1976),Borowsky (1976),Shier , Witzgall (1978), Kennedy , Gentle (1978) , Wolfe (1979),Porter Winstanley (1979),Barr , Affleck - Graves , Money,Hart (1980a), Harter (1981), Madsen (1985),Gonin, du - Toit (1987), Fichet (1987b)

درحالت جواب دستگاه معادلات فوق تشخیص نرم L_∞ روشهای مشابه وجـود دارد (برای اطلاعات بیشترعلاقه مندان می توانندبه مقالات منتـخب زیرومنابع مندرج درآنـــان رجوع نمایند:

Kelley (1958), Goldstein , Cheney (1958) , Cheney, Goldstein (1958) , Stiefel (1960), Veidinger (1960)

حداقل مربعات مستقیم وغیرمستقیم (IDLS , DLS) رامیتوان بترتیب برای دستگاه‌های (۳۹) و (۴۰) بکاربرد. تابع ابژکتیونرم L_1 مشابه حداقل مربعات دومرحله‌ای برای برآورد α رامی توان به‌شکل زیرتعریف نمود.

$$\min_{\alpha}: \sum_{i=1}^{n} |y_i - P_i^T Z\alpha| \tag{۴۱}$$

که y_i عنصر i ام y و P_i^T ردیف i ام $P=(X^T X)^{-1} X^T$ می باشد(نگاه کنیدبه(Fair (1974)). (1982) Amemiya بـــا مقایسه مسئله (۴۱) باتعبیر Theil از 2SLS

$$\min_{\alpha}: \sum_{i=1}^{n} (y_i - P_i^T Z\alpha)^2 \tag{۴۲}$$

وتعبیر 2SLS بعنوان برآوردکننده متغیرابزاری ، به‌عبارت دیگر ،حداقل کردن،

$$\min_{\alpha}: \sum_{i=1}^{n} (P_i^T y - P_i^T Z\alpha)^2 \tag{۴۳}$$

حداقل قدرمطلق انحرافات دومرحله‌ای (2SLAD) رابه شکل زیرتعریف می نماید

$$\min_{\alpha}: \sum_{i=1}^{n} |P_i^T y - P_i^T Z\alpha| \tag{۴۴}$$

(1982) Amemiya دوایده راترکیب نموده و 2SLAD رابعنوان خانواده‌ای ازبرآوردکننده‌ها حاصل ازحداقل کردن تابع زیررا پیشنهادمینماید .

$$\min_{\alpha}: \sum_{i=1}^{n} |qy_i + (1-q)P_i^T y - P_i^T Z\alpha| \tag{۴۵}$$

که q پارامتری است که توسط محقق تعیین می شود. اگر $q=0$ باشد، مسئله (۴۵) معادل (۴۴)می شودومنتج به برآوردی می شودکه بصورت مجانبی معادل 2SLS می‌باشد. اگر $q=1$ باشد(۴۵) معادل (۴۱) خواهدبود. برای هرمقدار $q\in[0,\infty)$ (1982) Amemiya سازگـــاری‌قوی 2SLAD راثابت نموده واریا مجانبی آنراتحت سه حالت مختلف توزیع‌نرمال ،نرمال جزئی وغیرنرمال **u** و **v** ارائـــه می نماید. (1983) Powell نرمال بودن مجانبی برآوردکننده‌های پیشنهـــادی (1982) Amemiya رابرای توزیع جمله خطادرحالت خیلی کلی تربیـــان می دارد.

(1982) Amemiya همچنین گزینه LAD مشابه بـــا 2SLS را

پیشنهاد می‌نماید. یکبار IDLAD درمورد هرمعادله فرم خلاصه شده بکاربرده شده و $\pi^{\wedge}$ محاسبه می شود. سپس با حداقل کردن عبارت زیر،

$$\min_{\beta}: \sum_{i=1}^{n} |y_i - X_i{}^T\pi^{\wedge}\theta - X_{1i}{}^T\beta| \qquad (۴۶)$$

$\theta^{\wedge}$ و $\beta^{\wedge}$ بدست می آیند این برآوردکننده راوی حداقل قدرمطلق انحرافات دومرحله‌ای دوبله (D2SLAD) نامگذاری می‌نماید. بحث مشابه‌ای برای مقادیر مختلف q نیزارائه شده‌است. (1983) Powell قضیه‌ای مبنی برمعادل بودن جانبی زیرطبقه برآوردکننده های D2SLAD راارائه می دهد. این نتیجه مشابه معادل بودن نمونه محدودتعبیر Theil از 2SLS وتعبیروی ازمتغیرابزاری می باشد.

(1970) Hunt و Glahe بعنوان پیشقدمان معرفی نرم L_1 در دستگاه معادلات همزمان، خواص نمونه کوچك برآوردکننده‌های حداقل مربعات وقدرمطلق‌ها برای یك دستگاه دومعادله‌ای همزمان فوق تشخیص راازطریق آزمایشات مونت کارلومقایسه می نمایند. برآوردکننده‌هائی که استفاده شدندعبارتنداز DLAD ، DLS ، IDLAD ، IDLS ، 2SLAD ، 2SLS ،تمام مقایسات برای برآوردکننده‌های حداقل مربعات وقدرمطلق های دومرحله‌ای مستقیم وغیرمستقیم برای اندازه‌های مختلف نمونه ۱۰ و ۲۰ باملاحظه حالات مختلف همخطی، هترسکدستیسیتی وسوتشخیص انجام شدند. آنها نتیجه‌گیری نمودندکه برآوردکننده‌های نرم L_1 درالگوهایی که دارای ساختاری مشابه مطالعات آنان با اندازه نمونه خیلی کوچك وتوزیع تصادفی خطاها باشد برتریا لااقل هم ترازبرآوردکننده نرم L_2 است همین ساختاررا (1974) Glahe و Dowling , Hunt باتوزیع نرمال و لاپلاس خطاهابکارمی برندبرآوردکننده‌هادر مطالعات آنان DLAD و DLS و 2SLAD و 2SLS می باشند. آنها نتیجه‌گیری می‌نمایند که ۱۰۰٪ بهترین نتایج درحالت توزیع لاپلاس و ۳۷/۵٪ ازبهترین نتایج درحالت توزیع نرمال رابرآوردکننده‌های نرم L_1 به خود تخصیص‌دادند.

(1977) Westlund و Nyquist مطالعه‌ای مشابه بایك دستگاه سه معادله‌ای فوق تشخیص همزمان باجملات خطاکه دارای توزیع های باثبات قرینه باشند را انجام می دهند. برآوردکننده‌های مورداستفاده این مطالعه مشابه (1970) Hunt

و Glahe که دربا لاذکرشدمی باشد. آنهانتیجه‌گیری می کنندکه باتوزیع نرمال، برآوردکننده‌های نرم L_2 مطلوب هستند. درحالت غیرنرمال برآوردکننده‌های نرم L_1 باافزایش درجه غیرنرمال بودن عملکردبهتری رادارند. زمانیکه اندازه نمونه‌افزایش می یابد، عملکردنسبی 2SLS و DLS نیزافزایش می یابد. درحالت توزیع نرمال 2SLS بهترین است وبرای توزیع‌های غیرنرمال 2SLAD بهترین گزینه بوده و IDLAD خیلی نزدیک آن است وبرای حالت غیرنرمال IDLAD خیلی نیرومندتراز 2SLAD بنظرمی رسد.

۵- جنبه هـــای آمـــاری

ازآنجائیکه معیارنرم L_1 بسط های جالبی درآماریافته است، دراین قسمت نگاهی گذرابربعضی ازوجوه آن درزمینه های مختلف آمارخواهیم داشت.

۱۰۵- توزیع‌نمونه گیـــری

Meyer , Glauber(1964). Rice, white (1964),Ashar,Wallace (1963) (1973),Smith ,Hall (1972) , Fama , Roll (1971) ,Glahe , Hunt(1970) Holland (1977), Ramsay (1977) , Brecht (1976) , Kiountouzis Pfaffenberger , Dinkel (1978), Rosenberg, Carlson (1977), Hill

خواص نمونه کوچك براز ش نرم L_1 راازطریق روش مونت کارلودرشرایط مختلف آزموده‌اندکارآئی نسبی این برآوردکننده به حداقل مربعات درزمانی که توزیع خطاها دارای دم های کلفت باشدتائیدگردیده‌است.

Wilson (1978) نتیجه گیری می نمایدکه زمانیکه خطاهاازتوزیع نرمال آلوده (Contaminted) تبعیت می کنندبرآوردکننده‌نرم L_1 ۸۰٪ کاراتر ازحداقل مربعات می باشد. زمانیکه نقاط دورافتاده حضوردارند، برآوردکننده نرم L_1 کاراترمی شود. دیدگاه وی نیزمونت کارلواست ودامنه وسیعی ازآزمایشات رادربرمیگیرد.

Cogger (1979) باانجام‌پیش بینی های ex - post از الگوهـای سری زمانی اتورگرسیونرم L_1 و L_2 راباهم مقایسه می نماید. این مقایسه توجه‌بیشتر به دیدگاه نرم L_1 دربرآورد الگوهای ARIMA (moving average

Integrated autoregressive) برای داده های سری زمانی رادرکارهای عملی توصیه می نماید.

برای رگرسیون چندمتغیره باتوزیع قرینه جمله خطا، (1977) Carlson و Rosenberg نشان دادندکه خطاهادربرآوردنرم L_1 تقریبا" دارای توزیع نرمال با میانگین صفروماتریس واریانس کواریانس $\delta^2(X^TX)^{-1}$ است که δ^2/n واریانس میانه‌خطاها می باشد(همچنین نگاه کنیدبه (1984) Tvejte و Sposito (1984) Ronner).آنها نتیجه‌می گیرندکه برآوردهای نرم L_1 دررگرسیون هائی که توزیع‌های خطادارای کورتوسیس بلندباشدواریانس‌کمتری از حداقل مربعات دارند(همچنین نگاه کنیدبه (1983) Steiger و Bloomfield).

(1973) Hartley و Sielken و(1985) Farebrother نشان داده‌اندکه اگرتوزیع خطاها قرینه باشدوبرآوردهای نرم L_1 یکتاباشند، مسئله را می توان طوری فرموله کردکه برآوردهای نااریب داشته باشند. بحثی مشابه درحالت کلی نرم L_p رامیتوان در(1982) Sposito پیداکرد.

(1978) Koenker و Bassett نشان میدهندکه برآوردهای نرم L_1 پارامترهای رگرسیون درالگوی عمومی سازگاروبطورمجانبی دارای توزیع Gauss با ماتریس کواریانس $\delta^2(X^TX)^{-1}$ می باشد δ^2/n واریانس مجانبی میانه‌نمونه ازنمونه های تصادفی به‌اندازه n ازتوزیع خطاگرفته شده‌است (نگاه کنیدبه، (1982) Koenker و Bassett و (1984) Bassett و Koenker و (1983) Bloomfield , Steiger و(1982) Oberhofer و(1988) wu).

یك روش تقریب ساده برای محاسبه اریب وچولگی برآوردهای نرم L_1 توسط (1987) Withers ارائه می شودکه نشان می دهدکه اریب وچولگی $\hat{\beta}$ متناسب باگشتاور سوم متغیرهای مستقل است ، مسئله گشتاوردرنرم L_1 توسط (1965) Rice و Hobby موردبحث قرارمی گیرند.

(1987a,b) Dupacova ازابزارحساب دیفرانسیل ناپذیرو epi - Convergence برای پیداکردن خواص مجانبی برآوردنرم L_1 مقیداستفاده می نماید. خواص جالب مجانبی برآوردکننده Boscovich که حداقل کردن نرم L_1 خطاها باتوجه به قیدمیانگین صفرپسماندها می باشدرا میتوان در(1985) Bassett

و Koenker پیدانمود. برازش نرم L_1 برای الگوهای رگرسیون سانسورد (Censored) (یاتوبیت (Tobit) سانسورد) توسط (86 , 1984) Powell معرفی شده‌اند. (1984) Paarsch با استفاده از آزمایشات مونت کارلو نشان داد که برآوردکننده Powell نه صحیح و نه باثبات است .

(1979) Steiger و Gross یک مشابه نرم L_1 از برآوردکننده نرم L_2 برای پارامترهای اتورگرسیون های مرتبه محدود و ایستا را بکار می برند. این برآوردکننده نشان داده است که قویاً " سازگار می باشد. استناد آنان به آزمایشات مونت کارلو است (همچنین برای بحث بیشتر نگاه کنید به (1983) Steiger و Bloomfield)

۲۰۵ـ استنباط آماری

توزیع مجانبی سه آماره نرم L_1 (آزمون های Wald ،نسبت درستنمائی وضریب لاگرانژ) فرضیه‌های خطی برای الگوی خطی عمومی توسط (1982) Bassett و Koenker بحث می شود. آنها توزیع مجانبی را برای خانواده بزرگی از توزیع ها بدست می آورند. نشان داده شده است که این آزمون ها در شرایط منظم بودن سست توزیع خطا و طرح ،رفتارکای ـ مربع حدی یکسانی دارند. مقایسه این آزمون ها بر مبنای آزمایشات مونت کارلو توسط (1987) Koenker ارائه میشود. از آنجائیکه برآوردکننده نرم L_1 بطور مجانبی از توزیع نرمال تبعیت می کند، (1987) Narula و Stangenhaus با استفاده از روش مونت کارلو اندازه نمونه را تعیین می کنند که تقریب توزیع نرمال را می توان برای ساختن فواصل اعتماد و آزمون فرضیه برروی پارامترهای رگرسیون نرم L_1 بکاربرد. روشهای مقایسه برای استودنت کردن میانه نمونه که قابل گسترش به رگرسیون نرم L_1 می باشد توسط (1984) Sheather , Mckean بحث می شود و براساس آن ،آزمون و فواصل اعتماد توسط (1987) Mckean و Sheather مقایسه می شوند.

دوضریب تعیین برای رگرسیون نرم L_1 در (1987) Sievers , Mckean آمده است. یک خانواده از آزمون ها برای هتروسکدستیسیتی بر مبنای چندک های رگرسیون توسط (1982 b) Bassett , Koenker معرفی می شوند. مطالعات

اخیردراستنباط آماری وآنالیزواریانس رامیتوان درمقالات زیریافت :

(1977) Armstrong et al (1983)، Siegel (1976)،

Sheather (1987)، Mckean , Shrader (1987)، Mckean

Shrader (1987)، Strangenhaus و(1987) Hettmansperger

و Brown و (1987) Khan و Tracy و(1987) Vajda

و(1987) Sheather و(1987) Fedorov . برای بررسی سایرویژگی ها

رجوع کنیدبه(1987a) Fichet و(1987) LeCalve .

۳۰۵ـ آمارچندمتغیره

درروش خوشه بندی معمول ،متریك یافاصله اقلیدسی بعنوان یك تابع مقدارحقیقی مناسب برای ساختن معیارنامتشابهی استفاده‌می شود۰(نگاه کنیدبه (1983 a) Bidabad). Spath (1976) متریك L_1 رابعنوان معیاری برای مسئله خوشه بندی استفاده نمود۰ تعدیل وبسط بیشتررا میتوان در (1987) Spath پیداکرد۰(1987) Kaufman , Rousseeuw یك گزینه‌نوع‌نرم L_1 که در روش K - medoid استفاده میشودرامعرفی نمودکه متوسط نامتشابهی تمام اقلام مجموعه داده هاراازنزدیکترین medoid حداقل می نماید۰(1987) Trauwaert و (1987) Jajuga متریك L_1 رادرروش خوشه بندی fuzzy برای ISODATA (Iterative Self Organizing Data Analysis Technique) بکارمی برند۰ (1987) Trauwaert نشان می دهدکه درزمان حضورنقاط دورافتاده یاخطای داده‌ها ،متریك L_1 برترازفاصله L_2 میباشد۰

یك نگرش نرم L_1 مشابه مقیاس بندی چندبعدی را(1988) Heiser (همچنین نگاه کنیدبه ،(1980) Critchley) وتجزیه تحلیل های ارتباط (Correspondence)را(1987) Heiser معرفی می نمایند۰

تجزیه تحلیل های تمایزنرم L_p نیرومندتوسط (1984) Haussler و(1985a) Hawkins بحث می شود۰برآوردنرم L_1 مولفه‌های اصلی توسط (1987) Watson و Galpin موردملاحظه قرارگرفته‌است ۰

۴-۵- برآوردچگالی غیرپارامتری

نرم L_1 همچنین درآمار غیرپارامتری وبرآوردچگالی مورداستفاده واقع شده است . روش برآوردچگالی ازطریق تابع هسته Parzen انجام می شود.(1976a,b, c) Abou - Jaouda و Wanger (1979, 80) و Devroye شرایط همگرائی نرم L_1 برآوردهای چگالی هسته راارائه می‌نمایند.(1983, 85) Devroye ویژگی های کامل سازگاری نرم L_1 برآوردچگالی Rosenblatt - Parzen راارائه می دهد. Devroye نتیجه‌گیری می نمایدکه تمام انواع سازگاری های نرم L_1 معادل هستند. (1987) Gyorfi سازگاری نرم L_1 برای برآوردهای چگالی هیستوگرام وهسته درنمونه های یکنواخت وبشدت مختلط راثابت می کند. (1985) Gyorfi ، Devroye برآوردغیرپارامتری نرم L_1 رابطورکامل تشریح می نمایند.قضیه حدمرکزی نرم L_p برای برآوردکننده‌های هسته‌ای چگالی ونرمال بودن مجانبی آنان درشرایط مختلف برآوردکننده‌های ساده وزنی وغیروزنی نرم L_p وسانسورکردن تصادفی توسط (1987 , 88) Csorgo , Horvath و (1987) Horvath و (1987) Csorgo , Gombay , Horvath بحث می شوند.انتخاب bandwidth دربرآوردرگرسیون غیرپارامتری را(1987) Marron نشان می دهد.باذکریک مثال وی نتیجه می گیردکه این یک مسئله برازش است.(1987) Welsh برآوردکننده هسته ای نرم L_1 ساده تابع تنکی (Sparsity) راملاحظه نموده وخواص مجانبی آنرابررسی می نماید.معیارهای Cross - Validation نرم های L_1 و L_2 برای خانواده بزرگی از برآوردکننده‌های هسته ای توسط (1987, 88) Rossi , Brunk مطالعه می شود.(1987) Gyorfi , Vander Meulen خواص همگرائی چگالی - آزادبرآوردکننده‌های آنتروپی Shannon رابررسی نموده وسازگاری نرم L_1 آنان راثابت می نمایند(1987) Perez , Fernandez Palacin Munoz برآوردتابع چندك راباستفاده چندجمله ای های Berstein ملاحظه نموده ورفتارنمونه بزرگ آنرادرنرم L_1 می آزمایند.برای مقایسه برآورد کننده‌های نرم L_1 و L_2 پارامترهای Weibull نگاه کنیدبه (1981)

Shier , Lawrence وبرای دیدگاه غیرپارامتری رگرسیون چندك بـه (1988) Sarda و Lejeune رجوع نمائید.

۵.۵ آمـــارنیرومنـــد

یکی ازمهمترین خواص روشهای نرم L_1 مقاومت به نقاط پرت یاخودرومی باشد. این خاصیت آنرایکی ازمهمترین تکنیك های آمارنیرومندساخته است. (1987) Huber اشاره می نمایدکه نرم L_1 دردوزمینه اصلی برآوردنیرومندایفای نقش می نماید. میانه نمونه نقش مهمی رادرآمارنیرومنددارد. میانه نمونه ساده ترین مثالی است که ازحداقـــل کردن نرم L_1 انحرافات حاصل می شود، بنابراین، نرم L_1 بزرگترین اریب مجانبی را حداقل می کند که میتواندباآلودگی (Contamination) نامتقارن ایجادشده باشد. بنابراین، درحالاتی که کنترل اریب مهمترازواریانس برآورداست برآوردنیرومنـــد مسئله می باشد. ثانیا " ، روش نرم L_1 ساده ترین برآوردکننده بانقطه شکنندگی بـالای موجوداست. بنابراین می تواندنقطه شروع خوبی برای برآوردکننده‌های تکراری باشدکه‌اگر بایك جواب اولیه بدشروع شوند به‌جوابهای بی معنی می رسنـــد واز آنجـــائیکـــــه مقاوم به نقاط پرت است میتواندبعنوان شروع خوبی برای برچیدن نقاط خودروبکاربـــرده شود(همچنین نگاه کنیدبه (1974) Taylor و(1977) Holland , Welsch (1977 , 78) Harvey و(1980) Armstrong ,Frome , Sklar و(1986) Antoch et al و(1987) Antoch و(1987) و(1988b) Bassett و Portnoy).

این تکنیك برای رگرسیون چندجمله‌ای باآزمونی برای درجه چندجمله‌ای برای چندك هـــای رگرسیون توسط (1983 , 84) Jureckova و(1984) Sen و Jureckova ملاحظه گردیده‌است. همین حالت برای رگرسیون غیرخطی توســـــط (1988) Prochazka طرحمیشود. (1987) Ronchetti مفاهیم اصلی آمارنیرومنـد رابرمبنای تابع‌نفوذ وهمچنین درارتباط بانرم L_1 مرورمی نماید(همچنین نگاه‌کنیدبـــه (1986) Galpin). برای الگوریتم های محاسباتی دررگرسیون بانفوذ محصوربـــه (1988) Marazzi رجوع شود. (1974) Ekblom خوب بودن آمـــاری روش های مختلف درزمانی که درموردمسئله رگرسیون بکاربرده میشوند راازطریـــــق

وی رابطه Ekblom (1987) آزمایشات مونت کارلو مورد بحث قرار می دهد و در برآورد نرم L_1 بعنوان حالت حدی یک نرم L_p یا برآورد هیوبر (Huber) را نشان می دهد (1984) Haussler و (1985a) Watson مسئله آنالیز تمایز نرم L_p نیرومند را ملاحظه مینماید. برآوردهای نیرومند مولفه های اصلی (نگاه کنید به Bidabad (1983c)) برمبنای فرمولاسیون نرم L_1 توسط (1987) Galpin , Hawkins مورد بحث قرار می گیرد. خواص ریسک توزیعی مجانبی برآوردکننده های نرم L_1 پیش آزمون و کوچکسازی (Shrinkage) توسط (1987) Sen , Saleh بررسی می شود. برآوردکننده نرم L_1 همچنین عضوی از برآورد کننده های M و R میباشد (برای بحث بیشتر نگاه کنید به (1983) Steiger و Bloomfield).

۶- کاربرد

روش های نرم L_1 بطور گسترده در رشته های مختلف علم توسعه یافته است و بعنوان یک ابزار تحلیل قوی در تجزیه تحلیل رفتار انسانی و پدیده های طبیعی بکار می رود. شاخه های متعددی از علوم در ریاضیات کاربردی، آمار و تحلیل داده ها چون اقتصادسنجی، زیست سنجی، روان سنجی، جامعه سنجی، فن سنجی، تحقیق در عملیات، مدیریت، فیزیک، شیمی، نجوم، طب، صنعت، مهندسی، جغرافیا و نظائر آن بمقدار زیادی وابسته به این روش هستند.

فرض توزیع نرمال خطاها همیشه برای متغیرهای اقتصادی همانند سایر متغیرها و داده ها برقرار نمی باشد و بنابراین همیشه با واریانس محدود سروکار نداریم. یک واریانس نا محدود به معنی توزیع دم کلفت خطاها با تعداد زیادی نقاط پرت می باشد. از آنجائیکه حداقل مربعات وزن زیادی به نقاط پرت می دهد بشدت وابسته به نمونه می شود. بنابراین در این حالت حداقل مربعات برآوردکننده ضعیفی می باشد. البته، توزیع متغیرهای اقتصادی یا اجتماعی هیچزمان واریانس بی نهایت را نشان نمی دهند. به هرحال، همانطورکه توسط Mandelbrot (1961, 63) بحث می شود، موضوع اصلی این نیست که گشتاور دوم توزیع عملا" بینهایت باشد، ولی دامنه بین دهک ها در رابطه با دامنه بین چارک ها به اندازه کافی بزرگ است که فرد را توجیه نماید که واریانس بی نهایت است. بدین ترتیب برآوردکننده ای همانند برآوردکننده نرم L_1 که وزن کمتری به نقاط پرت

می دهدبوضوح مرجح است .

توزیع درآمدشخصی اززمان پارتو ۱۸۹۶ دارای این مشخصه شناخته شده است . (1972) Ganger , Orr مستنداتی درخصوصیات سری زمانی متغیرهای اقتصادی که این خاصیت رادارندارائه می نماید. خیلی دیگرمتغیرهای اقتصادی مانندعایدی اوراق بهادار ،قیمت های احتکازی ،قیمت کالاوسهام ،اشتغال ،اندازه‌های سرمایه‌بنگاه‌های تجاری ،معادلات تقاضا ،نرخ بهره ،جریان های نقدی خزانه ،بیمه وانتظارات قیمت همگی درون طبقه توزیع خطاباواریانس بی نهایت قرارمی گیرند(نگاه کنیدبه (Quandt 1981) و Goldfeld (1977) و Westlund و Nyquist (1965) و Fama (1971) و Sharpe).

(1959) Hoffenberg و Arrow نرم L_1 رادرزمینه تقاضای بین صنایع استفاده نمودند. (1964) Glauber و Meyer نرم های L_1 و L_2 رامستقیما " مقایسه می کنند. آنها الگوهای سرمایه‌گذاری خودرابرای یک نمونه توسط هر دوبرآوردکننده برآوردنموده وباپیش بینی نمونه ex - post آنهارامی آزمایند. آنها نتیجه گیری می نمایندکه بجزمعدودی استثنا ،برآوردنرم L_1 برنرم L_2 فائق آمده، حتی بامعیارهائی نظیرمجموع مربع خطاهای پیش بینی که بطورمعمول برای حداقل مربعات حداقل بنظرمیرسد. (1971) Sharpe برآوردکننده‌های نرم L_1 و L_2 را برای اوراق بهاداروسهام مقایسه می نمایند. بحث مشابهی دربودجه بندی سرمایه توسط (1978) Dietrich , Cornell ارائه می گردد. (Carter) و Money و Affleck - Graves همچنین تحقیق رابابکارگیری نرم L_p وتاکیدبرعوامل موثربربرآوردضرائب یک الگوی سهام فردی انجام می دهند . (1987)

Kaergard برآوردکننده‌های نرم L_1 L_2 , L_∞ برای سرمایه‌گذاری دانمارک ازطریق قدرت پیش بینی سالهای زوج ازبرآوردانجام شده بروی سالهای فردبرای یک دوره طولانی رامقایسه می کند(1988) Hattenschwiler روش برنامه‌ریزی هدف رادررابطه باتوابع برازش نرم L_1 برروی چندالگوی‌برنامه‌ریزی خطی تجزیه شده برای سیاست تامین موادغذائی سوئیس استفاده‌می نماید(برای توضیح ارتباط برنامه‌ریزی هدف نگاه کنیدبه (1984 a Bidabad). سایرکاربردهای توابع برازشی نرم L_1 درالگوهای برنامه‌ریزی خودکفائی محصولات غذائی ،سهمیه بندی

غذائی والگوی تغییروتنظیم علوفه همگی توسط (1988) Hattenschwiler ذکروبحث می شوند.

(1979) Wilson رگرسیون نرم L_1 رابرای برآوردهزینه آماری دربعد حمل ونقل استفاده نمود. (1966) Chisman برآوردکننده نرم L_1 رابرای تعیین زمانهای استانداردمشاغل که درآن عناصرکاراساساً"برای تمام مشاغل باستثنای کیفیت هرنوع عنصرکاری استفاده شده درمیان مشاغل متفاوت باشدبکارمی برد. (1977) Armstrong , Frome به این برآوردکننده برای مولفه سیکل روندیك سری زمانی اقتصادی اشاره می نماید.

(1955) Charnes , Cooper , Ferguson برآوردبهینه دستمزدانجام کارکارکنان راباحل مسئله نرم L_1 ازطریق برنامه‌ریزی خطی ارائه می نمایند. کاربرد نرم L_1 درنظریه مکان یابی ازاهمیت ویژه‌ای برخورداراست زیرابا این متریك فاصله قائم دونقطه رادردستگاه دکارتی دوبعدی بخوبی میتوان ملاحظه کرد(نگاه کنیدبه(1970) Cabot et al و(1971,72) Wesolowsky , Love و (1978) Drezner , Wesolowsky و(1978) Ratliff , Picard و (1979) Morris , Verdini و(1983) Megiddo , Tamir و (1987) Calamai , Conn ،همچنین نگاه کنیدبه کتاب شناسی (1984) Drext و Domschke).(1987a) Farebrother نرم L_1 را درموردنظریه تصمیم‌گروه بکارمی برد. (1987) Mitchell نرم L_1 رابرای پیداکردن کوتاه ترین مسیربرای حرکت‌یك ربات ازمیان موانع‌بکارمی برد. نرم L_1 در شیمی توسط (1978) Fausett , Weber ،درژئوفیزیك توسط (1966) Smith , Dougherty ،(1973) Claerbout ,Muir و(1979) Taylor , Banks , McCoy درنجوم توسط (1987) Rousseeuw درپروسه های فیزیکی وفارماکوکینتیك توسط (1980) Frome , Yakatan و(1987 a) Gonin , Money بکارگرفته شده‌است ، برای بیان مکانیکی نرم L_1 نگاه کنیدبه (1987d) Farebrother. کاربرد نرم L_1 درسیستم‌های قدرت‌برای برآوردوضعیت استاتیك توسط (1982)

Anderson (1965) ارائه می شود. Kotiuga , Vidyasagar
استفاده برآوردنرم L_1 رابرای تامین ضرائب غیرمنفی درمعادلات زمان خطی پیشنهاد
می نماید. برای کاربرددرداده های اندازه گیری مداری نگاه کنیدبه (et al (1968
Mudrov
.

۷- شقـــوق دیگـــــر

(1977a) Narula , Wellington حداقل کردن مجموع وزنی قدرمطلق
خطاهاراپیشنهادمی نمایند. به عبارت دیگرحداقل کردن عبارت $\Sigma w_i |u_i|$. یک
الگوریتم برای این مسئله پیشنهادمی شود. (1977b) Wellington
و Narula یک حالت خاص ازفرمولاسیون فوق رابانام حداقل مجموع نسبی خطاها
عنوان می کنند. دراین مسئله w_i مساوی $1/|y_i|$ قرارداده می شوند(همچنین
نگاه کنیدبه تذکر (1980) Steiger , Bloomfield).
(1977c) Narula , Wellington الگوریتمی برای رگرسیون نرم
L_1 زمانی که مقیدبه عبورازمیانگین متغیرهامی باشدارائه می نمایند(همچنین به تذکر
(1987c) Farebrother نگاه کنید). درموردبرآوردنرم L_1
الگوریتم های مقیدالگوریتم هائی توسط (1971) Young و(1977) Hultz
و Armstrong و(1977 , 78) Barrodale , Roberts و(1980a,b)
Bartels , Conn (1980) Armstrong , Kung پیشنهاد
شده است . یك الگوریتم برای رگرسیون نرم L_1 بامتغیرهای مجازی توسط (1977)
Armstrong , Frome ارائه می شود. (1986) Womersley
یك الگوریتم گرادیان ساده شده رابرای رگرسیون نرم L_1 سانسورشده معرفی می نماید.
درزمینه رگرسیون گام به گام وانتخاب متغیرنیزالگوریتم های خاص درحالت نرم L_1
وجوددارند(نگاه کنیدبه (1974) Roodman (1977) Hansen
و Gentle (1979 , 83) Narula , Wellington و(1981) Narula
و Wellington و(1981) Dinkel , Pfaffenberger (1982a)
Armstrong , Kung).

یک الگوریتم برای چندک های رگرسیون بوسیله (1984) Wellington و Narula بیان شده است . محاسبه بهترین رگرسیون یک ـ طرف نرم L_1 که پیدا کردن یک تابع تقریب است که همیشه پائین یا بالای تابع باشد توسط (1970) Lewis ارائه گردیده است . برای روشهای عددی که برآوردهائی را پیدا می کند که مرز بالای مطلق انحرافات را حداقل می کنند نگاه کنید به (1987) Gaivoronski .

(1981) Arthanari , Dodge ترکیب محدبی از توابع ابژکتیو نرمهای L_1 و L_2 را برای پیدا کردن برآورد جدیدی برای الگوی رگرسیون خطی پیشنهاد می کنند . (1984) Dodge این روش را به ترکیب محدب توابع ابژکتیو برآوردکننده M هیوبر (Huber) و نرم L_1 بسط میدهد (1987) Jureckova و Dodge نشان دادند که ترکیب محدب موردبحث برآوردهای نرم L_1 و L_2 را میتوان طوری انطباق داد که برآوردکننده سازگاری از واریانس مجانبی برآوردکننده ایجاد شده جدید را حداقل نماید . در (1988) Dodge , Jureckova بحث می شود که ترکیب انطباقی برآوردکننده‌های M و نرم L_1 را میتوان بطور بهینه طوری انتخاب کرد که حداقل واریانس مجانبی ممکن را بدست آورد .

بجای حداقل کردن قدر مطلق انحرافات ، (1988) Nyquist قدر مطلق انحرافات عمودی از خط رگرسیون را حداقل می نماید . در این مقاله جنبه های محاسباتی این برآوردکننده ملاحظه می شود و ربط آن به مسئله تعقیب تصویر برای برآورد پراکندگی چند متغیره روشن می گردد . (1987) Spath , Watson نیز روش تقریب خطی نرم L_1 عمودی را معرفی می نمایند . کاربرد معیار فاصله عمودی برای نرم L_2 و نرم کلی L_p را میتوان در (1982, 86b) Spath و (1982b) Watson و (1983) Wulff یافت .

(1984) Rousseeuw یک روش جدید برآورد بنام رگرسیون " حداقل میانه مربعات " را پیشنهاد می نماید . این برآوردکننده با حداقل کردن عبارت $med (u_i^2)$ با توجه به β بدست می آید . برآوردکننده حاصل میتواند در مقابل اثر تقریبا " ۵۰٪ آلودگی در داده ها مقاوم باشد . یک کتاب کاربردی در این زمینه توسط (1987) Leroy و Rousseeuw نوشته شده است . الگوریتم های محاسباتی این برآوردکننده

رامیتوان در (1987) Steele , Souvaine (1986) Steiger

و Steele پیدا نمود.

زمانیکه تعداد مشاهدات در مقایسه با تعداد مجهولات زیاد میباشد، باید بهتر باشد که مشاهدات را به چند خوشه مجهول تجزیه نمود و به دنبال بردارهای رگرسیون مربوطه طوری گشت که میانگین مجموع نرم L_p بردار پسماند حداقل باشد. این ترکیب خوشه بندی و رگرسیون، رگرسیون خوشه وار (Clusterwise) نامیده می شود، مطالعه موردی و مقایسه عددی برای رگرسیون های نرم L_1 خطی خوشه وار در (1986a) Spath آمده است. برای الگوریتم های رگرسیون نرم L_1 خطی خوشه وار به (1986c) Spath و (1987) Meier و برای توضیح رگرسیون خوشه وار نگاه کنید به (1985, 87) Spath.

نرم L_1 در جداول یك طرفه و دوطرفه توسط (1976b,79) Frome و Armstrong و (1981) Kvanli, Buckley بکاربرده شده است (همچنین برای بحث های عمومی در این مورد نگاه کنید به (1983) Steiger و Bloomfield).

کاربردهای دیگری از نرم L_1 را در آمار U توسط (1987) Chun ، در نگرشی بیزی (Baysian) توسط (1987) Cap, Militky و رگرسیون ایزوتونی توسط (1987) Salvador , Menendez و تخصیص نمونه توسط (1987) Sadasivan , Melaku و روش میانگین ها توسط (1987) Kventon یافت.

فصل سوم

" معرفی الگوریتم های جدید "

۱- مقـــــــدمـــــــه

درفصل قبل ،جنبه های مختلف رگرسیون نرم L_1 را مرورکردیم وکارهای انجام شده در این زمینه ذکرگردیدند. ملاحظه کردیم که معیارنرم L_1 جای خودرادرتحلیل های علمی پیداکرده است. ازآنجائیکه ازلحاظ محاسباتی باسایرمعیارهاهمانندنرم L_2 قابل مقایسه نیست ، احتیاج به کاربیشتری داردتاآن یك ابزارعملی شود. فرم بسته جواب برآوردکننده نرم L_1 هنوزبدست نیامده است ، وبنابراین ، استنباط های بیشتر ازخواص این برآورد کننده رامشکل می نماید. هرتلاشی برای ارائه الگوریتم های محاسباتی کاراکه بتواند ویژگی های مختلف مسئله رانیزروشن تر نمایدمطلوب است. ازاین بابت دراین فصل سعی براین خواهیم داشت که یك روش عمومی برای حل مسئله رگرسیون خطی نرم L_1 ارائه نمائیم. الگوریتم های پیشنهادی برمبنای یك روش خاص نزولی بوده وتکنیك مشتق گسسته رااستفاده می نمایند. طرح های اولیه این الگوریتم هاتوسط (Bidabad (1987a, b, 88a,b بحث شده اند. باتوسعه این الگوریتم ها ،گونه های کاراتری معرفی می شوندکه نشان خواهیم دادکه عملکردبهتری ازسایرالگوریتم های موجوددارند. الگوی رگرسیون زیررا درنظربگیرید ،

$$y_i = \sum_{j=1}^{m} \beta_j x_{ji} + u_i \qquad i=1,...,n \tag{۱}$$

که $\beta j,\ j=1,...,m$ پارامترهای جامعه می باشندکه بایدبرآوردشوند ، y_i و x_{ji} و u_i به ترتیب متغیرهای تابع، مستقل وتصادفی هستند. جهت سهولت دراین فصل x_{ji} بجای x_{ij} استفاده می شوند. می خواهیم βj ها راباحداقل کردن عبارت زیر برآورد کنیم ،

$$S = \sum_{i=1}^{n} |u_i\hat{}| = \sum_{i=1}^{n} |y_i - \sum_{j=1}^{m} \beta_j\hat{} x_{ji}| \tag{۲}$$

فرض کنید $m=1$ ، بنابراین عبارت(۲) به شکل زیرخلاصه می شود ،

$$S = \sum_{i=1}^{n} |y_i - y_i\hat{}| = \sum_{i=1}^{n} |y_i - \beta\hat{} x_i| = \sum_{i=1}^{n} S_i \tag{۳}$$

یك عنصرنمونه ، $S_i = |y_i - \beta\hat{} x_i|$ رامیتوان بعنوان یك خط شکسته در صفحه $S x \beta\hat{}$ متشکل ازدونیم خط درنظرگرفت. S_i مقدارحداقل خودراکه مساوی صفر است درمقدارزیرپیداخواهدکرد.

$$\hat{\beta}_i = y_i / x_i \qquad (۴)$$

شیب این نیم خط ها در دست چپ و راست $\hat{\beta}_i$ بترتیب برابر $-|x_i|$ و $|x_i|$ می باشند. بنابراین S_i ها همگی محدب هستند و نتیجتا " مجموع آنها S نیز محدب است و شیب آن در هر مقدار $\hat{\beta}$ برابر با مجموع شیب های S_i ها در آن مقدار $\hat{\beta}$ می باشد از آنجائیکه شیب هر S_i فقط در $\hat{\beta}_i$ مربوطه تغییر می کند، حداقل S در یکی از $\hat{\beta}_i$ ها واقع خواهد شد. بدین ترتیب، خط رگرسیون از نقطه مبداء و با شیبی مساوی $\hat{\beta}_i$ که S را حداقل می کند می گذرد. به عبارت دیگر، برای پیدا کردن برآوردنرم L_1 β فقط احتیاج به پیدا کردن یك مشاهده داریم (نگاه کنید به Karst (1958) و Taylor(1974)) در این رابطه (1974) Taylor نتیجه گیری می نماید که " این البته بدین معنی است که خط رگرسیون باید از مشاهده ای مرتبط با حداقل کردن i عبور کند. بنابراین، خط رگرسیون با نقطه مبدا و مشاهده مرتبط با حداقل کردن $\hat{\beta}_i$ مشخص می گردد". ولی وی این دیدگاه حداقل کردن S با توجه به زیرنویس i را دنبال نمی کند. در این فصل همانند Bidabad (1987a,88a) اولین تلاش توسعه این نقطه نظر است. در قسمت بعد، بعد از نوشتن مجدد (۳) به یك شکل مناسب، مقدار i با استفاده از مشتق گسسته تعیین می شود.

با بحثی مشابه میتوان نتیجه گرفت که هرگاه تعداد پارامترها m باشد، m مشاهده باید روی فوق صفحه رگرسیون قرار گیرند. به عبارت دیگر، m معادله به شکل (۵) لازمند تا فوق صفحه رگرسیون را مشخص کنند (همچنین نگاه کنید به فصل اول).

$$y_i - \sum_{j=1}^{m} \hat{\beta}_j x_{ji} = 0 \qquad (۵)$$

حال با ساده ترین الگوی رگرسیون بحث را دنبال می کنیم.

۲- رگرسیون خطی ساده باقید

برای الگوی (۱) حالت یك متغیر مستقل بدون عرض از مبداء به شکل $y_i = \beta x_i + u_i$ را در نظر بگیرید. برای پیدا کردن برآوردنرم L_1 β، روش زیر پیشنهاد می شود.

$$S = \sum_{i=1}^{n} |u_i| = \sum_{i=1}^{n} |y_i - \beta x_i| = \sum_{i=1}^{n} (y_i - \beta x_i)\,\mathrm{sgn}(y_i - \beta x_i) = \sum_{i=1}^{n} |x_i|\,(y_i/x_i - \beta)\,\mathrm{sgn}(y_i/x_i - \beta) \qquad (۶)$$

فرض کنید $z_i = y_i / x_i$ و z_i رابترتیب نزولی مرتب می کنیم. $z_i, \ i=1,\ldots,n$
مرتب شده‌حاصل‌را $z_h, \ h=1,\ldots,n$ نامگذاری می کنیم. عناصر z_h باید خاصیت زیررادارا باشند،

$$z_h > z_l \quad \text{if } h < l \qquad h,l=1,\ldots,n.$$

(۶) رابا مشاهدات مرتب شده می نویسم،

$$S = \sum_{h=1}^{n} |x_h|(z_h - \beta)\,\mathrm{sgn}(z_h-\beta) \tag{۷}$$

مشاهده‌ای که روی خط رگرسیون قرارمی گیردرا (x_{t+1}, y_{t+1}) نامگذاری می کنیم که مشاهده $(t+1)$ ام درآرایه z_h می باشد. مقدار z_h برابرشیب یك شعاع است که مبداء شروع شده ازمشاهده h ام عبورمی کند. بنابراین ،

اگر	$h < t+1$	پس	$z_h > \beta$	و	$u_h > 0$
اگر	$h = t+1$	پس	$z_h = \beta$	و	$u_h = 0$
اگر	$h > t+1$	پس	$z_h < \beta$	و	$u_h < 0$

حال (۷) رامیتوانیم به شکل زیربنویسیم

$$S = \sum_{h=1}^{t} |x_h|(z_h - \beta) - \sum_{h=t+1}^{n} |x_h|(z_h - \beta) \tag{۸}$$

برای پیداکردن حداقل S ،چون β و t هردومجهول هستنداحتیاج به مشتق گرفتن ازآن نسبت به β وزیرنویس t داریم. توجه کنیدکه β دارای دامنه‌پیوسته و t دارای دامنه گسسته است . بدین ترتیب،

$$\frac{\delta S}{\delta \beta} = -\sum_{h=1}^{t} |x_h| + \sum_{h=t+1}^{n} |x_h| \tag{۹}$$

مشتق S باتوجه به زیرنویس t بایدمشتق گسسته باشد(نگاه کنیدبه (1978 Bender, Orszag , Clarke(1983)) :

$$\frac{\mathrm{Delta}(S)}{\mathrm{Delta}(t)} = \lim_{\mathrm{Delta}(t)\to 1} \frac{S[t+\mathrm{Delta}(t)]-S(t)}{\mathrm{Delta}(t)} = S(t+1)-S(t) =$$

$$\sum_{h=1}^{t+1} |x_h|(z_h-\beta) - \sum_{h=t+2}^{n} |x_h|(z_h-\beta) -$$

$$\sum_{h=1}^{t} |x_h|(z_h-\beta) + \sum_{h=t+1}^{n} |x_h|(z_h-\beta) = 0$$

یا ،

$$2|x_{t+1}|(z_{t+1}-\beta) = 0$$

بدین ترتیب ،

$$\hat{\beta} = z_{t+1} = y_{t+1}/x_{t+1} \tag{۱۰}$$

بایدتوجه کردکه اگر $x_{t+1}=0$ باشدمقدار $z_{t+1}=\infty$ است وزمانیکه آرایه z مرتب می شود، مقادیربینهایت z درابتداوانتهای آرایه واقع می شوندو سبب مسئله‌ای در(۱۰) نخواهندشد. بخاطرداشته باشیدکه‌معادله (۱۰) همانندرابطه (۴) است. معادلات (۹) و(۱۰) دومعادله بادومجهول t, β هستند. متغیر t رابابازنویسی (۹) به شکل زیرمیتوان یافت ،

$$D_k \equiv \sum_{h=1}^{k}|x_h| - \sum_{h=k+1}^{n}|x_h| \qquad k=1,\ldots,n \tag{۱۱}$$

واضح است که D_k یك تابع افزایشی از k می باشد. هرگاه k ازیك به n افزایش می یابد، D_k مقادیرمختلفی راحاصل می کندکه ازمنفی به مثبت افزایش می یابند. بنابراین ، مقدمتا " k مساوی یك قراردادەمیشودو D_k برمبنای آن محاسبه می گردد. اگر D_k منفی است ، k رابه مقداریك واحدافزایش میدهیم وروش راادامه می دهیم تا D_k مثبت شود. زمانیکه D_k به‌اولین مقدارمثبت رسید، $t+1=k$ می باشد. بااین روش مقدار $t+1$ پیدامی شود. مشاهده مرتبط بااین زیرنویس $(t+1=k)$ را انتخاب می کنیم (x_{t+1}, y_{t+1}) برآوردنرم L_1 β باجایگزینی مقادیر y_{t+1} و x_{t+1} در(۱۰) بدست خواهدآمد(نگاه کنیدبه (Bidabad (1987a, 88a).

روش ارائه شده اساسا " روش میانه وزنی (Laplace (1818 می باشدکه‌در فصل قبل موردبحث قرارگرفت . به‌عبارت دیگراین جواب قدمهائی است که لاپلاس برداشت. اختلاف اصلی دربدست آوردن ریاضی جواب است . لاپلاس این جواب راباتحلیل جبری ویژگی های تابع ابژکتیونرم L_1 بدست آورد، درصورتیکه روش فوق دقیقا " ازتابع ابژکتیومشتق می گیرد. نوآوری دیگراین روش کاربردمشتق گسسته جزئی برروی زیرنویس است که همراه‌با

مشتق معمولی برروی متغیرهای بادامنه پیوسته بکارگرفته شده است . بایدخاطرنشان ساخت که (1757) Boscovich این مسئله رابایك روش هندسی و(1958) Karst بایك روش تحلیلی حل نمود .

۲-۱- محاسبه میانه وزنی

دوسری محاسبات برای محاسبهٔ میانه وزنی لازم است . یك الگوریتم مرتب کننده برای مرتب کردن آرایه z وذخیره زیرنویس های مربوطه برای محاسبه قسمت دوم عملیات جهت پیداکردن مقداربهینه k بااستفاده D_k که در (۱۱) تعریف شده است .

الگوریتم های مرتب کننده کارآئی برای اولین قسمت محاسبات وجوددارند . الگوریتم های "quicksort" متعلق به (1961,62) Hoare و "quickersort" متعلق به (1965) Scowen و "Sort" متعلق به (1969) Singleton ازعملکردکارآئی خوبی برخوردارهستند . برای قسمت دوم محاسبات یك روش خاص وجودندارد، ولی (1980a) Steiger , Bloomfield روش "Partial sorting" متعلق به (1971) Chambers برای ارائه یك روش کارآبرای ترکیب دو مرحله مرتب کردن آرایه z وپیداکردن مقدار k رااستفاده‌می نمایند . برتری‌این روش درمرتب کردن قسمت های کوچکتری ازآرایه z می باشدتاتمام عناصرآن . باتعدیل هائی ، این روش رابرای الگوریتم های پیشنهادی درقسمت های بعدی بکارخواهیم گرفت . این روش رامیتوان به شکل تابع زیربیان نمود .

```
FUNCTION LWMED (n,ys,w,l)
Step 0) Initialization.
        Real: ys(n), w(n).
        Integer: l(n), hi.
        Set: ii=0, shi=0, slo=0, sz=0, sp=0, sn=0.
Step 1) Compute left, middle and right sum of weights.
        Do loop for i=1,n: w(i)=|w(i)|; if ys(i)<0, then
        sn=sn+w(i), if ys(i)>0, then sp=sp+w(i), if ys(i)=0
```

then sz=sz+w(i); end do.

If shi≤slo then go to step 2.b, otherwise go to step 2.a.

Step 2) Assign subscripts for arrays.

a. Let: shi=0.

Do loop for i=1,n: if ys(i)≤0 go to continue, otherwise ii=ii+1, l(ii)=i, continue, end do.

Go to step 2.c.

b. Let: slo=0.

Do loop for i=1,n: if ys(i)>0 go to continue, otherwise ii=ii+1, l(ii)=i, continue, end do.

c. Let: lo=1, hi=ii.

Step 3) Check for solution.

If hi>lo+1 then go to step 4, otherwise lwmed=l(lo). If lo=hi return, otherwise if ys(l(lo))≤ys(l(hi)) go to step 3.a, otherwise lt=l(lo), l(lo)=l(hi), l(hi)=lt, lwmed=l(lo).

a. If shi+w(l(hi))>slo+w(l(lo)) then set lwmed=l(hi), otherwise return.

Step 4) Divide the string into two halves then sort.

Set: mid=(lo+hi)/2, lop=lo+1, lt=l(mid), l(mid)=l(lop), l(lop)=lt.

a. If ys(l(lop))≤ys(l(hi)) then go to step 4.b, otherwise lt=l(lop), l(lop)=l(hi), l(hi)=lt.

b. If ys(l(lo))≤ys(l(hi)) then go to step 4.c, otherwise lt=l(lo), l(lo)=l(hi), l(hi)=lt.

c. If ys(l(lop))≤ys(l(lo)) then go to step 5, otherwise lt=l(lop), l(lop)=l(lo), l(lo)=lt.

Step 5) Compute the accumulation of weights.

Let: lwmed=l(lo), i=lop, j=hi, xt=ys(lwmed), tlo=slo, thi=shi.

a. Set: tlo=tlo+w(l(i)), i=i+1.

If ys(l(i))<xt then go to step 5.a, otherwise go to step 5.b.

b. Let: thi=thi+w(l(j)), j=j-1.

If ys(l(j))>xt then go to step 5.b, otherwise if j≤i

```
        then go to step 6, otherwise lt=l(i), l(i)=l(j),
        l(j)=lt, go to step 5.a.
Step 6) Test for solution.
        Let: test=w(lwmed).
        If i≠j then go to step 6.a, otherwise
        test=test+w(l(i)), i=i+1, j=j-1.
     a. If test≥|thi-tlo| then return, otherwise, if tlo>thi
        then step 6.b, otherwise slo=tlo+test, lo=i, go to
        step 3.
     b. Let: shi=thi+test, lo=lop, hi=j.
        Go to step 3.
END
```

۲-۲- بحث برالگوی m پارامتری

روشی که برای الگوی یك پارامتری ساده بکاربرده شدرانمی توان به سهولت به الگوی m پارامتری تعمیم داد. این‌شکل بدلیل این مسئلهٔ است که نمی‌توانیم مشاهدات راطوری مرتب کنیم که پسماندهای آنان بطورصعودی یانزولی مرتب باشند. به عبارت دیگر، برای اعمال کردن تکنیك مشتق گسسته، (۲) بایدمقدمتا " به شکل زیربازنویسی شود :

$$S = \sum_{h=1}^{t} (y_h - \sum_{j=1}^{m} \beta_j x_{jh}) - \sum_{h=t+1}^{n} (y_h - \sum_{j=1}^{m} \beta_j x_{jh}) \qquad (۱۲)$$

عبارت (۱۲) که عاری ازعلامت قدرمطلق است تعمیم (۸) به m پارامتراست. ولی اولین مشکل پیداکردن " منطقی" است که ماراقادرسازدتا (۱۲) راتشکیل دهیم. بعبارت دیگر، احتیاج داریم که‌مشاهدات راطوری مرتب کنیم که زمانیکه h کمتر، مساوی ویا بیشتراز t+1 می باشد، u_h بترتیب بزرگتر، مساوی ویاکمترازصفرباشد، باافزایش h ازیك به n پسماندمربوطه u_h کاهش می یابد.

اگر(۲) رامیتوانستیم به شکل (۱۲) بنویسیم، مجددا " می توانستیم ازآن نسبت به β_j برای تمام $j=1,...,m$ و t مشتق بگیریم. مشتق S نسبت به β_j برای $j=1,...,m$ برابرخواهدبود

$$\frac{\delta S}{\delta \beta_j} = -\sum_{h=1}^{t} x_{jh} + \sum_{h=t+1}^{n} x_{jh} \qquad j=1,...,m \qquad (۱۳)$$

برای مشتق گرفتن از S نسبت به متغیردامنه گسسته t خواهیم داشت :

$$\frac{\mathrm{Delta}(S)}{\mathrm{Delta}(t)} = \lim_{\mathrm{Delta}(t)\to 1} \frac{S[t+\mathrm{Delta}(t)]-S(t)}{\mathrm{Delta}(t)} = S(t+1)-S(t) =$$

$$\sum_{h=1}^{t+1}\left(y_h - \sum_{j=1}^{m}\beta_j x_{jh}\right) - \sum_{h=t+2}^{n}\left(y_h - \sum_{j=1}^{m}\beta_j x_{jh}\right) -$$

$$\sum_{h=1}^{t+1}\left(y_h - \sum_{j=1}^{m}\beta_j x_{jh}\right) + \sum_{h=t+1}^{n}\left(y_h - \sum_{j=1}^{m}\beta_j x_{jh}\right) =$$

$$2\left(y_{t+1} - \sum_{j=1}^{m}\beta_j x_{j,t+1}\right) = 0 \qquad (۱۴)$$

m مقدارمتمایزبرای t بااستفاده ازمقدار D_{kj} زیربرای هرمتغیرتوضیحـی رامی توان از (۱۳) محاسبه نمود. روش محاسبه t مشابه آن است که در (۱۱) برای الگوی یك پارامتری انجام دادیم .

$$D_{kj} \equiv \sum_{h=1}^{k} x_{jh} - \sum_{h=k+1}^{n} x_{jh} \qquad j=1,\ldots,m;\ i=1,\ldots,n \qquad (15)$$

بایدخاطرنشان ساخت که این m مقداربرای t ،همگی یك مقداربرای S در (۱۲) بدست خواهنددادزیرا تمام این خطاها مرتبط با این t هادارای مقادیرصفرهستند. این امر ازخاصیت پسماندهای صفررگرسیون نرم L_1 که درفصل اول بحث شدنیزروشن می باشدکه

m نقطه برروی فوق صفحه رگرسیون وجوددارندکه دارای پسماندهای صفرهستند. در (۱۲)، این m نقطه پشت سرهم قرارمی گیرند،زیرا مشاهدات راطوری مرتب کرده بودیم کههرگاه h ازیك به n افزایش می یابد،خطاها ازمقداربزرگترین مثبت به کوچکترین منفی کاهش خواهندیافت . بنابراین m خطای صفریکی پس ازدیگری تقریبا" دروسط دنباله n عضوی (۱۲) واقع می شوند. مطابق این m مقدار t ، mمشاهده شناسائی می شوند. باجایگزینی مقادیراین m مشاهده در (۱۴) ، m معادله پیدامی شودکه بطورهمزمان میتوان برای m مجهول β_j حل گردد. این m معادله آخرخاصیت رگرسیون نرم L_1 که فوقا" ذکرگردیدراتائیدمی نمایند.

ازآنجائیکه برای رگرسیون m پارامتری مرتب کردن پسماندهاقبل ازمحاسبه‌مقادیر بهینه β_j هاممکن نیست ،بایدروش دیگری برای محاسبه برآوردنرم L_1 پارامتری های رگرسیون خطی چندمتغیره ابداع نمود.

۳- رگرسیـون خطـی ساده بدون قیـد

برای پیداکردن برآوردهای نرم L_1 $\{\beta_j, j=1,...,m\}$ برای الگـوی (۱) سعی براین است که الگوریتم هائی برای جستجوی آن نقاطی که برروی صفحه رگرسیـون بهینه قرارمیگیرندپیشنهادشود. برای $m=2$ و $x_{1i}=0$ و $i=1,...,n$ (۱) به رگرسیون خطی ساده باقیدبافقط یك پارامترخلاصه می شود. برآوردنرم L_1 شیب β_2 برای الگوی ساده رامیتوان بااستفاده ازالگوریتم ارائه شده درقسمت های قبل بدست آورد. حال الگوی خطی ساده بدون قیدراکه درآن $m=2$ و $x_{1i}=1$ بـرای $i=1,...,n$ است راملاحظه کنید، به عبارت دیگر،

$$y_i = \beta_1 + \beta_2 x_{2i} + u_i \tag{۱۶}$$

۳-۱- الگوریتـم ۱

برای الگوی (۱۶)، تابع ابژکتیو S که بایدحداقل شودبرابراست با :

$$S = \sum_{i=1}^{n} |y_i - \beta_1 - \beta_2 x_{2i}| \tag{۱۷}$$

فرض کنید $k1$ زیرنویس باشدکه بدامنه‌یك تا n تعلق دارد، وفرض کنیدکـه مشاهده‌ام $k1$ ام (y_{k1}, x_{2k1}) نامزدبرای واقع شدن روی خط رگرسیون باشد. اگر اینطورباشد، $u_{k1}=0$ خواهدبودومیتوانیم مبدا مختصات YxX_2 رابدون هیچ اشکالی بنقطه (y_{k1}, x_{2k1}) منتقل کنیم. برای این کارباید تمام مشاهدات رابـه عنوان انحرافهائی ازنقطه (y_{k1}, x_{2k1}) بنویسیم

$$\begin{aligned} y_i^{k1} &= y_i - y_{k1} \qquad & i=1,...,n \\ x_{2i}^{k1} &= x_{2i} - x_{2k1} \qquad & i=1,...,n \end{aligned} \tag{۱۸}$$

که جملات بابا لانویس $k1$ نشانگرمقادیرمنحرف هستند. جملات راجابجا می کنیم،

$$\begin{aligned} y_i &= y_i^{k1} + y_{k1} \qquad & i=1,...,n \\ x_{2i} &= x_{2i}^{k1} + x_{2k1} \qquad & i=1,...,n \end{aligned} \tag{۱۹}$$

حال (۱۹) رادر (۱۷) جایگزین می‌نمائیم، مسئله حداقل سازی نرم L_1 رامیتوان بـه شکل زیرمجددا " تعریف نمود.

$$\min_{\beta_1, \beta_2}: S_{k1} = \sum_{i=1}^{n} |y_i^{k1} - \beta_2 x_{2i}^{k1} + (y_{k1} - \beta_1 - \beta_2 x_{2k1})| \tag{۲۰}$$

ازآنجائیکه فرض کردیم که مشاهده $k1$ ام روی خـــط رگـــرسیـــون اســـت .

$y_{k1}-\beta_1-\beta_2 x_{2k1}=0$ خواهدبود ،پس بدین ترتیب (۲۰) خلاصه می شودبه ،

$$\min_{\beta_2}: S_{k1} = \sum_{i=1}^{n} |y_i^{k1} - \beta_2 x_{2i}^{k1}| \tag{۲۱}$$

حل مسئله بهینه‌سازی (۲۱) همان چیزی است که برای الگوی خطی یك پارامتری بیان شد ،توجه نمائیدکه زمانیکه $u_{k1}=0$ ،مقدارحداقل (۲۱) مساوی مقدارحداقل (۱۷) می باشد .

β_2 حاصل ازحداقل کردن (۲۱) را β_2^{k1} می نمائیم ،باتغییر $k1$ ازیك به n وحداقل کردن (۲۱) ، $\beta_2^1, \ldots, \beta_2^n$ بدست می آیند ،حال مسئله این است : کدام مقدار $k1$ (۱۷) راحداقل می کند؟ به‌عبارت دیگرکدام مشاهدات برروی خط رگرسیون بهینه قراردارند؟ توجه کنیدکه درالگوی خطی دوپارامتری نرم L_1 حداقل دومشاهده‌با خطاهای صفروجوددارد .انتقال دستگاه مختصات YxX_2 به هرکدام ازاین دونقطه حداقل (۱۷) راتغییرنمی دهد .فرض کنیدنقاط p ام و q ام آن نقاطی هستندکه‌بر روی رگرسیون قراردارند ،بدین ترتیب ،

$$\begin{aligned} u_p &= y_p - \beta_1 - \beta_2 x_{2p} = 0 \\ u_q &= y_q - \beta_1 - \beta_2 x_{2q} = 0 \end{aligned} \tag{۲۲}$$

(۱۸) و (۱۹) رابرای $k1=p,q$ بازنویسی کرده وآنهارادر (۲۱) جایگزین می کنیم ،نتیجه برابراست با ،

$$S_p = \sum_{i=1}^{n} |y_i^p - \beta_2 x_{2i}^p| \tag{۲۳}$$

$$S_q = \sum_{i=1}^{n} |y_i^q - \beta_2 x_{2i}^q| \tag{۲۴}$$

بااستفاده‌از (۱۸) وبازنویسی (۲۳) و (۲۴) نشان میدهیم که S_p مساوی S_q است

$$S_p = \sum_{i=1}^{n} |y_i-\beta_2 x_{2i} - (y_p-\beta_2 x_{2p})| \tag{۲۵}$$

$$S_q = \sum_{i=1}^{n} |y_i-\beta_2 x_{2i} - (y_q-\beta_2 x_{2q})| \tag{۲۶}$$

S_p زمانی مساوی S_q است که اگروفقط اگردوپرانتزداخل علائم قدرمطلق در (۲۵) و (۲۶) مساوی باشند .این موضوع رامیتوان ازحل دومعادله (۲۲) برای β_1 بدست آورد .بدین معنی که ،

$$\beta_1 = y_p - \beta_2 x_{2p} = y_q - \beta_2 x_{2q} \tag{۲۷}$$

تساوی S_q و S_p بدین علت است که نقاط p و q برروی خط رگرسیون هستندو بنابراین $u_p=u_q=0$ است ・بدین ترتیب $S_p=S_q$ می باشدوبنابراین، مقادیر $\beta_2{}^p$ و $\beta_2{}^q$ حاصل ازحداقل‌کردن S_p یا S_q بایدمساوی باشند・این امریک معیاری برای پیداکردن β_2 مطلوب ازکلیه $\beta_2{}^{k1}$ هابرای $k1=1,\ldots,n$ راارائه می نماید・به این معنی که هرگاه $\beta_2{}^p=\beta_2{}^q$ باشد・مقداربرآورد β_2 رابا $\hat{\beta}_2$ مشخص می کنیم ・ مقدار $\hat{\beta}_1$ به سهولت از (۲۷) محاسبه می شود・حال تمام مراحل برای پیداکردن‌مقادیر $\hat{\beta}_1$ و $\hat{\beta}_2$ برای لگوی (۱۶) راخلاصه‌می کنیم・

الگوریتم خام ۱

قدم صفر) قراردهید $k1=1$ ・

قدم اول) (۱۸) رامحاسبه کنید・

قدم دوم) (۲۱) رابااستفاده روش میانه وزنی حداقل‌کرده و $\beta_2{}^{k1}$ راحساب کنید・

قدم سوم) آزمایش کنیداگر $\beta_2{}^{k1}=\beta_2{}^{k1-k}$ برای $0<k<k1$ قرار دهید $\hat{\beta}_2=\beta_2{}^{k1}$ و $\hat{\beta}_1=y_{k1}-\hat{\beta}_2 x_{2k1}$ وتوقف کنید،

قدم چهارم) $k1$ رابه میزان یك واحدافزایش داده وبه‌قدم اول بروید・

۳-۲- الگوریتم ۲

الگوریتم خام ۱ برای‌پیداکردن $\hat{\beta}_1$ و $\hat{\beta}_2$ ازلحاظ محاسباتی کارآنیست زیرا معمولا" نیازبه آزمون اکثرمشاهداتدارد・بنابراین برای اینکه این الگوریتم راکاراتر سازیم تغییراتی لازم می باشد・ بجای قراردادن $k1=1$ ، $k1$ رامساوی یک مقدار صحیح دلخواه a قرارمی دهیم a یک عددصحیح ازیک تا n است ・ حال فرض کنیدکه $u_a=y_a-\beta_1-\beta_2 x_{2a}=0$ است و (۲۱) رابه شکل زیربازنویسی می کنیم・

$$\min_{\beta_2}: \; S_a = \sum_{i=1}^{n} |y_i{}^a - \beta_2 x_{2i}{}^a| \qquad (۲۸)$$

حداقل‌کردن (۲۸) مقدار $\beta_2{}^a$ رامی دهدکه برابراست با

$$\beta_2{}^a = \frac{y_b{}^a}{x_{2b}{}^a} = \frac{y_b - y_a}{x_{2b} - x_{2a}} \qquad (۲۹)$$

به عبارت دیگر ،باچرخش روی مشاهده a ام ،نقطه دیگری همانند b پیدامی شود

که b اشاره به آن مشاهده‌ای دارد که دارای خطای صفر در جواب حداقل (۲۸) می باشد، بنابراین $u_b = y_b - \beta_1 - \beta_2 x_{2b} = 0$ حداقل S_a در (۲۸) را با S_a^* مشخص می کنیم ،پس ،

$$S_a^* = \sum_{i=1}^{n} |y_i^a - \beta_2^a x_{2i}^a| \qquad (۳۰)$$

توجه کنید

$$\beta_1 = y_a - \beta_2^a x_{2a} = y_b - \beta_2^a x_{2b} \qquad (۳۱)$$

این رابطه از ضرب (۲۹) در مخرج آن و جابجائی جملات حاصل می شود. (۳۱) را در (۱۷) جایگزین می کنیم.

$$S = \sum_{i=1}^{n} |y_i - y_a - \beta_2^a (x_{2i} - x_{2a})| = \sum_{i=1}^{n} |y_i - y_b - \beta_2^a (x_{2i} - x_{2a})|$$

یا

$$S = \sum_{i=1}^{n} |y_i^a - \beta_2^a x_{2i}^a| = \sum_{i=1}^{n} |y_i^b - \beta_2^a x_{2i}^b| \qquad (۳۲)$$

با استفاده از (۳۰) اولین مجموع در (۳۲) S_a^* است و مجموع دوم S_b است که در $\beta_2 = \beta_2^a$ ارزیابی شده. بدین ترتیب ،میتوان نتیجه گرفت که،

$$S_a^* = S_b \Big|_{\beta_2 = \beta_2^a} \qquad (۳۳)$$

S_a^* درحداقل است ولی $S_b \big|_{\beta_2 = \beta_2^a}$ هنوز میتواند برای سایر مقادیر β_2 حداقل شود. بنابراین نتیجه مهمی بدست می آید که.

$$S_a^* > S_b^* \qquad (۳۴)$$

نامساوی (۳۴) تضمین می کند که اگر یک نقطه دلخواه را برای انتقال مبداء مختصات به آن انتخاب کنیم و تابع ابژکتیو (۲۱) را حداقل کنیم نقطه دیگری پیدا می شود ،که انتقال مبدا مختصات به این نقطه جدیداً " پیدا شده کل مجموع مطلق خطاها را کاهش خواهد داد. بنابراین ،درهر انتقال به حداقل S نزدیکتر می شویم. با بحثی مشابه ،این نتیجه را میتوان به شکل زیر تعمیم داد،

$$S_a^* > S_b^* > S_c^* > S_d^* \ldots \qquad (۳۵)$$

توجه کنید که a یک نقطه شروع دلخواه است. نقطه b با حداقل کردن S_a ، c با

حداقل کردن S_b و d باحداقل کردن S_c ونظائرآن بدست می آیند.

حال مسئله است که چه موقع مقدار حداقل S حاصل خواهدشد. فرض کنید $S^* = S_f^*$ ؛ با انتقال مبداء مختصات به نقطه f وحداقل کردن S_f ،مشاهده g ام بدست می آید. زمانیکه $S^* = S_f^*$ با حداقل کردن S_g ،مشاهده f ام مجددا" پیدا می شودزیرا $S_g^* = S_f^* = S^*$ ومشاهدات g ام و f ام هردوبرروی خط رگرسیون نرم L_1 قراردارند. این نتیجه یك معیاری برای توقف دراین روش می تواندباشد.بدین ترتیب،

$$S_a^* > S_b^* > S_c^* > S_d^* \ldots > S_f^* = S_g^* = S^* \qquad (۳۶)$$

بایدخاطرنشان کردکه اگرجواب حداقل S یکتانباشد،بدین معنی که S یك قسمت افقی داشته باشد،این روش هرگاه به اولین جواب بهینه برسدمتوقف خواهدشد.

حال تمام مراحل این الگوریتم را برای پیداکردن برآوردنرم L_1 β_1 و β_2 درالگوی ساده خطی (۱۶) بیان می کنیم.

الگوریتم ۲

قدم صفر) مقدارصحیح دلخواه a رابین یك تا n انتخاب کرده و $k1 = a$ قرار می دهیم.

قدم اول) (۱۸) رابا $k1 = a$ محاسبه می کنیم.

قدم دوم) (۲۱) رابا استفاده ازروش میانه وزنی حداقل کرده ومشاهده مناسب روی خط راپیدا می کنیم فرض کنیدنقطه b.

قدم سوم) (۱۸) رابا $k1 = b$ محاسبه می نمائیم.

قدم چهارم) (۲۱) راحداقل کرده ومشاهده روی خط راپیدامیکنیم،مشاهده c.

قدم پنجم) آزمایش کنیداگر $c=a$ باشد $\hat{\beta}_2 = y_{bc}/x_{2bc}$, $\hat{\beta}_1 = y_c - \hat{\beta}_2 x_{2c}$ وتوقف می کنیم.

قدم ششم) قرارمی دهیم $a=b$ وبه قدم اول می رویم.

ازلحاظ عملیاتی قدم های زیرعملا" برداشته می شوند.برنامه زیردرفصل چهارم درمقایسه باسایرالگوریتم هابرای الگوی رگرسیون خطی دوپارامتری استفاده خواهدشد.

```
PROGRAM BL1S
Step 0) Initialization.
        Parameter: n.
        Real: y(n), x2(n), z(n), w(n).
        Integer: l(n).
        Set: k1=arbitrary, k1r=0, k1s=0, iter=0.
        Read (y(i), x2(i), i=1,n)
Step 1) Compute weights and ratios.
        Do loop for i=1,k1-1: w(i)=x2(i)-x2(k1),
        z(i)=(y(i)-y(k1))/w(i), end do.
        Set: w(k1)=0, z(k1)=0.
        Do loop for i=k1+1,n: w(i)=x2(i)-x2(k1),
        z(i)=(y(i)-y(k1))/w(i) end do.
        Set: iter=iter+1.
Step 2) Compute weighted median.
        Let: lm=LWMED(n,z,w,l).
Step 3) Check for optimality.
        Set: k1s=k1r, k1r=k1.
        If lm=k1s then go to step 4, otherwise k1=lm.
        Go to step 1.
Step 4) Compute the solution.
        Let b2=z(lm), b1=y(k1)-b2*x2(k1).
        Print b1, b2, k1, lm, iter.
        stop.
END
```

۴ـ تعمیم به m پارامتر

حال روش فوق را به الگوی دوپارامتری با قید تعمیم می دهیم ،به عبارتی ،

$$y_i = \beta_2 x_{2i} + \beta_3 x_{3i} + u_i \tag{۳۷}$$

قرار می دهیم

$$S = \sum_{i=1}^{n} |y_i - \beta_2 x_{2i} - \beta_3 x_{3i}| \tag{۳۸}$$

S رامیتوان به شکل زیرنوشت ،

$$\begin{aligned} S &= \sum_{i=1}^{n} |x_{2i}|\,|y_i/x_{2i} - \beta_2 - \beta_3 x_{3i}/x_{2i}| = \\ &\sum_{i=1}^{n} |x_{2i}|\,|y_i^{s1} - \beta_2 - \beta_3 x_{3i}^{s1}| \end{aligned} \tag{۳۹}$$

که

$$y_i^{s1} = y_i / x_{2i} \qquad i=1,...,n$$
$$x_{3i}^{s1} = x_{3i} / x_{2i} \qquad i=1,...,n \qquad (۴۰)$$

حداقل کردن (۳۹) مشابه روش گفته شده برای الگوی خطی ساده‌می باشدکه درقسمت قبل بیان گردید. یك تفاوت مهم برای حل (۳۹) درمقایسه با (۱۷) عبارت $|x_{2i}|$ است که در $|y_i^{s1}-\beta_2-\beta_3 x_{3i}^{s1}|$ ضرب شده‌است . این ضرب هنگام حداقل کردن (۳۹) هیچ مشکلی راایجادنمی نمایدزیرااگر y_i^{s1} و x_{3i}^{s1} راازنقطه $k1$ منحرف کنیم،

$$y_i^{s1k1} = y_i^{s1} - y_{k1}^{s1} \qquad i=1,...,n$$
$$x_{3i}^{s1k1} = x_{3i}^{s1} - x_{3k1}^{s1} \qquad i=1,...,n \qquad (۴۱)$$

میتوانیم (۳۹) رامشابه (۲۱) بازنویسی نمائیم بنابراین :

$$S_{k1} = \sum_{i=1}^{n} |x_{2i}| \, |y_i^{s1k1} - \beta_3 x_{3i}^{s1k1}| \qquad (۴۲)$$

برای حداقل کردن (۴۲) ،باید عبارت زیر رابکاربریم

$$S_{k1} = \sum_{i=1}^{n} |x_{2i} x_{3i}^{s1k1}| \, |y_i^{s1k1}/x_{3i}^{s1k1} - \beta_3| \qquad (۴۳)$$

براساس (1987a, 88a Bidabad) دربکارگیری مشتق گسسته در (۴۳) که درقسمت قبل توضیح داده‌شد ،جمله‌ای که دراولین علامت قدرمطلق قراردارد‌برای پیداکردن زیرنویس آن نقطه‌ای که برروی خط رگرسیون قرارمی گیردبکارگرفته‌می شود. در مقایسه باالگوی خطی بدون قیدساده (۱۶) این تفاوت اصلی می باشد.

درحالت ورودیك پارامترعرض ازمبداء درالگوی (۲۷) داریم ،

$$S = \sum_{i=1}^{n} |y_i - \beta_1 - \beta_2 x_{2i} - \beta_3 x_{3i}| \qquad (۴۴)$$

فرض کنید $k1$ یك زیرنویس دلخواه باشد ،بنابراین انتقال مبداء مختصات به‌نقطه $k1$ام باانحراف تمام مشاهدات ازاین نقطه بدست خواهدآمد ،به‌عبارت دیگر ،

$$y_i^{k1} = y_i - y_{k1} \qquad i=1,...,n$$
$$x_{2i}^{k1} = x_{2i} - x_{2k1} \qquad i=1,...,n \qquad (۴۵)$$
$$x_{3i}^{k1} = x_{3i} - x_{3k1} \qquad i=1,...,n$$

جملات (۴۵) راجابجانموده ودر (۴۴) جایگزین می کنیم ،پس داریم ،

$$S = \sum_{i=1}^{n} |y_i^{k1} - \beta_2 x_{2i}^{k1} - \beta_3 x_{3i}^{k1} + (y_{k1} - \beta_1 - \beta_2 x_{2k1} - \beta_3 x_{3k1})| \qquad (۴۶)$$

اگرمشاهده $k1$ام برروی صفحه رگرسیون است ، پس

$$y_{k1} - \beta_1 - \beta_2 x_{2k1} - \beta_3 x_{3k1} = 0 \qquad (۴۷)$$

بنابراین ،بجای حداقل کردن (۴۴) ،تابع زیرحداقل می شود

$$S_{k1} = \sum_{i=1}^{n} |y_i^{k1} - \beta_2 x_{2i}^{k1} - \beta_3 x_{3i}^{k1}| \qquad (۴۸)$$

حداقل کردن (۴۸) کاملا" مشابه (۳۸) بوده وبه‌شکل زیرعمل می شود

$$S_{k1} = \sum_{i=1}^{n} |x_{2i}^{k1}| |y_i^{k1}/x_{2i}^{k1} - \beta_2 - \beta_3 x_{3i}^{k1}/x_{2i}^{k1}| = \qquad (۴۹)$$

$$\sum_{i=1}^{n} |x_{2i}^{k1}| |y_i^{s1} - \beta_2 - \beta_3 x_{3i}^{s1}|$$

که

$$y_i^{s1} = y_i^{k1}/x_{2i}^{k1}$$
$$x_{3i}^{s1} = x_{3i}^{k1}/x_{2i}^{k1} \qquad (۵۰)$$

حال مجددا " مبدا ،مختصات دستگاه دوبعدی $Y^{s1}xX_3^{s1}$ رابه یك نقطه دلخواه $k2$ بامنحرف کردن y_i^{s1} و x_{3i}^{s1} از y_{k2}^{s1} و x_{3k2}^{s1} به شکل زیر منتقل می کنیم .

$$y_i^{s1k2} = y_i^{s1} - y_{k2}^{s1} \qquad i=1,...,n$$
$$x_{3i}^{s1k2} = x_{3i}^{s1} - x_{3k2}^{s1} \qquad i=1,...,n \qquad (۵۱)$$

باجابجائی جملات (۵۱) وجایگزینی آنهادر (۴۹) وبافرض اینکه نقطه $k2$ بر روی صفحه رگرسیون قراردارد ،می توانیم (۴۹) رابه شکل زیربنویسیم.

$$S_{k1k2} = \sum_{i=1}^{n} |x_{2i}^{k1}| |y_i^{s1k2} - \beta_3 x_{3i}^{s1k2}| \qquad (۵۲)$$

یا

$$S_{k1k2} = \sum_{i=1}^{n} |x_{2i}^{k1} x_{3i}^{s1k2}| |y_i^{s1k2}/x_{3i}^{s1k2} - \beta_3| \qquad (۵۳)$$

تابع ابژکتیو(۵۳) راهمانطوری که قبلا" گفته شدمیتوان حداقل کرد.حال ،این روش از (۴۹) تا (۵۳) میتواندبامقادیرمختلف $k2$ مانندالگوریتم ۲ که برای الگوی ساده خطی پیشنهادشدتکرارشود.زمانیکه آخرین نقطه (M) درفرآیندحداقل سازی (۵۳) پیدا می شود ،مبداء مختصات دستگاه سه‌بعدی YxX_2xX_3 ($k1$) بهاین نقطه‌جدیدا " پیداشده M منتقل می شودوتمام مراحل از (۴۵) تا (۵۳) مجددا " تکرارمی شودبااین استثناکه بجای منصوب کردن یك مقداردلخواه به $k2$, $k2$ رامساوی آخرین مقدار $k1$ قرارمی دهیم . این روش ادامه می یابدتانقطه پیداشده ازحداقل کردن (۵۳) مساوی مقدار قبلی $k1$ باشدمقادیر $\hat{\beta}_1$, $\hat{\beta}_2$, $\hat{\beta}_3$ رابرطبق فرمولهای زیرمی توان محاسبه نمود،

$$\begin{aligned} \hat{\beta}_3 &= y_M^{s1k2}/x_{3M}^{s1k2} \\ \hat{\beta}_2 &= y_{k2}^{s1} - \hat{\beta}_3 x_{3k2}^{s1} \\ \hat{\beta}_1 &= y_{k1} - \hat{\beta}_2 x_{2k1} - \hat{\beta}_3 x_{3k1} \end{aligned} \qquad (۵۴)$$

بایدخاطرنشان ساخت که می توان ثابت کردکه درهرقدم به حداقل S در(۴۴) نزدیک ـ ترمی شویم ۰اثبات همانندچیزی است که برای الگوی دوپارامتری (۱۶) ارائه‌گردید۰

میتوان نشان دادکه بازاء هردونقطه دلخواه $k1$ و $k2$ ،مقدار S_{k1k2} در(۵۳) برای هرجابجائی برای $k1$, $k2$ تغییرنخواهدکرد۰این موضوع رادرقسمت بعدنشان خواهیم داد۰

به هرحال

$$S_{k1k2} = S_{k2k1} \qquad (۵۵)$$

حال بدنبال روش اثبات (۲۸) تا (۳۶) می توانیم بنویسیم۰

$$S_{k1a}^* = S_{k1b}\Big|_{\beta_3=\beta_{3k1a}} \qquad (۵۶)$$

که S_{k1a}^* مقداربهینه S_{k1a} نسبت به β_3 در (۵۳) برای هرعددصحیح دلخواه a ازیک تا n رامشخص می نماید۰به‌عبارتی

$$S_{k1a}^* = \min_{\beta_3} (S_{k1a}) \qquad (۵۷)$$

زیرنویس b در(۵۶) مشخص کننده مشاهده‌ای است که باحداقل کردن S_{k1a} پیدا میشود ۰ ازآنجائیکه مشاهدات a ام و b ام هردوروی فوق صفحه رگرسیون هستند،با بحثی مشابه‌که از (۲۸) تا (۳۲) اظهارگردید،می توانیم نتیجه بگیریم که S_{k1a}^* مساوی S_{k1b} است که درآن مقداراز β_3 ارزیابی شده باشدکه باحداقل کردن S_{k1a} پیدا میشود۰ازآنجائیکه سمت چپ (۵۶) حداقل است وسمت راست آن راباسایرمقادیر β_3 میتوان کاهش داد،میتوان نتیجه گرفت که

$$S_{k1a}^* > S_{k1b}^* \qquad (۵۸)$$

برطبق (۵۵) می توانیم بنویسیم،

$$S_{k1a}^* > S_{k1b}^* = S_{bk1}^* \qquad (۵۹)$$

ازآنجائیکه درهرقدم یك مشاهده راازپایه حذف می کنیم وآنرابامشاهده‌جدیدا "پیدا

شده جایگزین می نمائیم ،داریم

$$S_{k1a}^* > S_{k1b}^* = S_{bk1}^* > S_{bc}^* = S_{cd}^* \qquad (۶۰)$$

یابطورکلی ،

$$S_{k1a}^* > S_{k1b}^* > S_{bc}^* > S_{cd}^* \qquad (۶۱)$$

جواب حداقل S^* زمانی بدست می آیدکه باورودمشاهده f ام به پایه‌نمی توانیم مقدارتابع ابژکتیوراکاهش دهیم ۰ یعنی مشاهده قبلی مجددا " بدست می آید ،به‌عبارت دیگر ،

$$S_{k1a}^* > S_{ab}^* > S_{bc}^* > S_{cd}^* > \ldots > S_{fg}^* = S_{gf}^* = S^* \qquad (۶۲)$$

رابطه (۶۲) تضمین می نمایدکه درهرقدم مابرروی رویه تابع ا بژکتیوپائین می آئیم۰

۴ـ۱ـ الگوریتم ۳

برای‌تعمیم الگوریتم فوق به الگوی m پارامتری خطی (۱) ،بایدتعدادپارامترها رابه‌همان روشی که درالگوی سه پارامتری توضیح داده شدکاهش دهیم ۰اگرالگوباقیداست ، می توانیم آنراباتقسیم تمام متغیرهای مستقل وتابع‌به یکی ازمتغیرهای مستقل به‌شکل زیر بدون قیدنمائیم۰

$$S = \sum_{i=1}^{n} |y_i - \sum_{j=2}^{m} \beta_j x_{ji}| = \sum_{i=1}^{n} |x_{2i}| \, |y_i / x_{2i} - \beta_2 - \sum_{j=3}^{m} \beta_j x_{ji}| \qquad (۶۳)$$

اگرالگوبدون قیداست ،باانحراف تمام مشاهدات ازیك مشاهده‌دلخواه میتوان آنراباقید کرد، به عبارتی

$$S = \sum_{i=1}^{n} |y_i - \beta_1 - \sum_{j=2}^{m} \beta_j x_{ji}| = \sum_{i=1}^{n} |y_i^{k1} - \sum_{j=2}^{m} \beta_j x_{ji}^{k1}| \qquad (۶۴)$$

که ،

$$y_i^{k1} = y_i - y_{k1} \qquad i=1,\ldots,n$$
$$x_{ji}^{k1} = x_{ji} - x_{jk1} \qquad i=1,\ldots,n;\ j=2,\ldots,m \qquad (۶۵)$$

بنابراین ،برطبق تبدیل های (۶۳) ، (۶۴) و (۶۵) ،هرالگوی m پارامتری میتواند به یك الگوی یك پارامتری باقیدساده شودوبنابراین همانندیك مسئله میانه وزنی حل شود۰

برای انجام این تبدیل ،اگر الگوبدون قیداست ،باید تمام مشاهدات را ازیك مشاهده دلخواه منحرف نمائیم والگورا باقیدکنیم ۰سپس تمام متغیرهای مستقل وتابع را به یك متغیر مستقل دلخواه تقسیم کنیم ۰ این امرسبب می شودکه الگوبدون قیدشود ۰دراین مرحله تعداد پارامترها را یکی کم کرده ایم۰ با ادامه این روش ،هرالگوی m پارامتری را می توانیم به یك الگوی یك پارامتری خلاصه کنیم۰ اگر الگوبا قیداست بایدبا تقسیم تمام متغیرهای مستقل شروع کنیم که سبب می شودکه الگوبدون قیدشود ۰حال ،می توانیم یکی ازپارامترها را با بکارگیری روش فوق برای تبدیل الگوی بدون قیدبه باقیدبکاربریم ۰ باحل الگوی یك پارامتری ،می توانیم سپس الگوی دوپارامتری وبعدازآن سه پارامتری ونظیرآن را حل کنیم ۰

ظریف ترین قسمت این الگوریتم این است که درنقطه شروع یکبار $k1,k2,...,k(m-1)$ بطوردلخواه انتخاب می شوندسپس این الگوریتم بهترین مقادیرممکن را به اعدادصحیح $k1,k2,...,k(m-1)$ منصوب می نماید ۰برای روشن کردن این روش فرض کنیدکه یك الگوی خطی بدون قیدچهارپارامتری داریم ۰ یکبار مقدار $k1$ بطوردلخواه انتخاب می شود ۰ تمام مشاهدات را ازمشاهده $k1$ منحرف می نمائیم ۰ ازاین طریق ،الگوی ما به یك الگوی باقیدسه پارامتری ساده می شود ۰با تقسیم تمام متغیرها به یکی ازمتغیرهای مستقل ،الگوی ما کاملا" شبیه (۴۴) می شود۰ باحل (۴۴) براساس الگوریتمی که قبلا" برای الگوی سه پارامتری توضیح داده شد ،زیرنویس M مرتبط بانقطه M ام بدست خواهدآمد ۰این نقطه جدیدا " پیداشده می باشدکه زیرنویس آن (M) به $k1$ منصوب می شود ۰مقدارقبلی $k1$ به $k2$ ومقدارقبلی $k2$ به $k3$ منصوب شده وتمام روش مجددا " تکرارمی شود ۰این روش زمانیکه M مساوی $k1$ است متوقف خواهدشد ۰

تکنیك انتخاب مهم فوق که برای چرخش برروی مبداء دستگاههای مختصات باابعاد مختلف اساسی است را می توان برای الگوهای باپارامترهای بیشترهمانندقبل بسط داد ۰ مجددا " بایدخاطرنشان ساخت که درهرمرحله به حداقل S در(۱) نزدیکترمی شویم ۰اثبات کاملا" مشابه الگوی سه پارامتری که فوقا " توضیح داده شده می باشد ۰تمام قدم های این روش ازقرارزیراست ،

الگوریتـــم ۳

قدم صفر) نقاط دلخواه $k1, ..., k(m-1)$ راانتخاب کنید.

قدم اول) قراردهید $counter = 1$

قدم دوم) تبدیل های (۶۳) تا (۶۸) رابرای ساده‌کردن الگوبه‌یك الگوی خطی باقیـــــد

تبدیل شده یك پارامتری انجام دهید.

قدم سوم) روش میانه وزنی رابرای پیداکردن نقطه M بکاربرید.

قدم چهارم) اگر M مساوی مقدارقبلی k (counter) هست وبه قدم پنجم بروید

درغیراینصورت قراردهید $k(\text{counter}) = M$ وبه قدم اول بروید.

قدم پنجم) اگر $\text{counter} = m-1$ به قدم ششم برویدودرغیراینصورت قـــرار

دهید $\text{counter} = \text{counter}+1$ وبه‌قدم چهارم بروید.

قدم ششم) قراردهید $k(m) = M$ بااستفاده مقادیرنهائی $k1, ..., k(m-1)$

دستگاه معادلات زیررابرای $\hat{\beta}_j$ برای $j=1, ..., m$ حل کنید.

$$\sum_{j=1}^{m} \hat{\beta}_j x_{jki} = y_{ki} \qquad i=1, ..., m$$

۴-۲- الگوریتـــم ۴

درالگوریتم ۳ برای حل هرالگوی m پارامتری ،الگوی ساده شده $m-1$ پارامتری بایداول حل شوداین موضوع سبب پرهزینه بودن این الگوریتم می شود.زیرابرای حل یـــك الگوی m پارامتری بایدمقدارتابع ابژکتیوالگوی $m-1$ پارامتری راتامقدارحداقل آن کاهش دهیم وسپس به الگوی m پارامتری برگردیم .درمراحل میـــانــی بــــرای $m-1, m-2, ..., 1$ این مراحل تکرار میشوند .برای کارآترساختن این الگوریتم‌تکنیـــك انتخاب دیگری راباید‌اتخاذکرد .یکبار $m-1$ مشاهده بطوردلخواه انتخاب می شوند تابع S بااستفاده از(۶۳) و (۶۴) همانندالگوریتم ۳ به یك مسئله یك پارامتـــری میانه وزنی ساده می شود .باحل مسئله میانه وزنی یك مشاهده جدید M پیدامی شود. حال نقطه جدیدا " پیداشده جایگزین قدیمی ترین مشاهده‌ای که‌واردپایه شـدمی شود.ایـــــن روش تاآنجاادامه می یابدکه نتوانیم مشاهده دیگری رابیرون ازمجموعه فعلی نقاط پایـــه ونقطه حذف شده دردورقبل بیابیم ،به عبارت دیگر ،دراین الگوریتم $m-1$ نقطه بعنـــوان

يك جواب شدنی پایه انتخاب می شوند. باچرخش برروی این جواب یك نقطه جدیدپیدا می شودکه بایدواردپایه شودوجایگزین نقطه دیگری شودکه قبلا" درپایه بود. این روش زمانیکه نتوانیم باورودهرنقطه دیگری مقدارتابع ابژکتیوراکاهش دهیم متوقف می شود. از این بابت این الگوریتم شباهتی به روش سیمپلکس برنامه ریزی خطی دارد.

الگوریتم ۴ همچنین یك روش نزولی است ،زیرادرهردورمقدار S کاهش می یابد. اثبات کاملا" شبیه موردالگوریتم ۳ است که قبلا" ذکرشد. اختلاف اصلی این الگوریتم با الگوریتم ۳ درانتخاب نقاطی است که بایدازپایه حذف گردند. بنابراین ،درهرقدم احتیاج نداریم که بعضی نقاط رابرای حل الگوهای بااندازه کوچکتردرپایه نگهداریم. به‌عبارت دیگر ،درهر m دورتمام m نقطه‌ای که درپایه هستنددورانداخته می شوندو m نقطه جدیدواردمی شوند. درالگوریتم ۳ بعضی نقاط راتارسیدن به یك نقطه بهینه برای الگوی ساده‌شده‌بااندازه‌کوچکتردرپایه نگه میداشتیم. بایدخاطرنشان ساخت که یك نقطه ممکن است ازپایه بیرون شودومجددا" درطی m دورواردپایه گردد. این امرمتناقض نیست و عملا" مخصوصا" زمانی‌که‌نزدیك جواب بهینه هستیم اتفاق می افتد.

همانطورکه درقسمت بعدی خواهیم دیدبدلیل خواص متعدداین الگوریتم تغییرات متعددی درمراحل محاسباتی این الگوریتم برای شروع محاسبات اتخاذخواهدشد. قبل از اینکه این الگوریتم راپیچیده‌ترنمائیم مناسب است که قدم های کلی آنرابشکل زیربرشماریم

الگوریتم ۴

قدم صفر) نقاط دلخواه $k1,\dots,k(m)$ راانتخاب کنید

قدم اول) تبدیل های (۶۳) تا (۶۸) رابرای ساده‌کردن الگوبه یك الگوی خطی باقید تبدیل شده یك پارمتری انجام دهید.

قدم دورم) روش میانه وزنی رابرای پیداکردن نقطه M بکاربرید.

قدم سوم) M رابه‌یکی از $k1,\dots,k(m-1)$ کمدرپایه ازهمه‌قدیمی تراست منصوب کنید.

قدم چهارم) آزمایش کنیداگرنقطه جدیدا" پیداشده M که واردپایه می شودنقطه‌ای راکه دردور $m-1$ قبل واردپایه شده‌بود باربیرون می نمایدبه

قدم پنجم بروید، درغیراینصورت به قدم اول بروید.

قدم پنجم) قراردهید $k(m)= M$ بااستفاده ازمقادیرنهائی $k1,\ldots,k(m)$

دستگاه معادلات زیررابرای β_j برای $j=1,\ldots,m$ حل کنید.

$$\sum_{j=1}^{m} \beta_j \hat{x}_{jki} = y_{ki} \qquad i=1,\ldots,m$$

وتوقف کنید.

بعدازبحث درخواص الگوریتم ۴، قدم های محاسباتی این الگوریتم رابطورتشریحی تر ارائه خواهیم کرد. بااستفادهازاین خواص میتوانیم این الگوریتم راخیلی کارآترنمائیم.

۴-۳- خواص

قبل ازبحث درباره خواص الگوریتم های پیشنهادی مخصوصا " الگوریتم ۴، تابع ابژکتیونرم L_1 الگوی چهارپارامتری زیررا ملاحظه نمائید.

$$S = \sum_{i=1}^{n} |y_i - \beta_0 - \beta_1 x_{1i} - \beta_2 x_{2i} - \beta_3 x_{3i}| \tag{۶۶}$$

مراحلی رابایدبرای ساده کردن (۶۶) به یك مسئله میانه وزنی باواردکردن سه عدد صحیح $k3, k2, k1$ بین یك تا n همانندزیربگذرانیم. توجه کنیدکه درتابع (۱) تغییرات کمی داده ایم تاعرض ازمبداء رابعنوان یك جمله ساده درمعادله (۶۶) واردنمائیم همانندتبدیل های (۶۳) تا (۶۵) قدم های زیربطورپیاپی برداشته می شود. تمام مشاهدات را ازمشاهده $k1$ ام منحرف می‌نمائیم.

$$\begin{aligned} y_i^{k1} &= y_i - y_{k1} && i=1,\ldots,n \\ x_{ji}^{k1} &= x_{ji} - x_{jk1} && i=1,\ldots,n;\ j=1,2,3 \end{aligned} \tag{۶۷}$$

تابع ابژکتیومی شود،

$$S_{k1} = \sum_{i=1}^{n} |y_i^{k1} - \beta_1 - \beta_2 x_{2i}^{k1} - \beta_3 x_{3i}^{k1}| \tag{۶۹}$$

y_i^{k1} و x_{ji}^{k1} رابرای $j=2,3$ بر x_{1i}^{k1} برای تمام $i=1,\ldots,n$ تقسیم می کنیم

$$y_i^{s1} = y_i^{k1}/x_{1i}^{k1} \qquad i=1,\ldots,n \tag{۷۰}$$

$$x_{ji}^{s1} = x_{ji}^{k1}/x_{1i}^{k1} \qquad i=1,\ldots,n;\ j=2,3 \tag{۷۱}$$

تابع ابژکتیو می شود ،

$$S_{k1} = \sum_{i=1}^{n} |x_{1i}^{k1}|\,|y_i^{s1}-\beta_1-\beta_2 x_{2i}^{s1}-\beta_3 x_{3i}^{s1}| \qquad (۷۲)$$

y_i و x_{ji}^{s1} را برای $j=2,3$ ازمشاهده $k2$ ام منحرف می کنیم،

$$y_i^{k2} = y_i^{s1} - y_{k2}^{s1} \qquad i=1,\ldots,n \qquad (۷۳)$$

$$x_{ji}^{k2} = x_{ji}^{s1} - x_{jk2}^{s1} \qquad i=1,\ldots,n,\ j=2,3 \qquad (۷۴)$$

تابع ابژکتیو ساده می شود به

$$S_{k1k2} = \sum_{i=1}^{n} |x_{1i}^{k1}|\,|y_i^{k2}-\beta_2 x_{2i}^{k2}-\beta_3 x_{3i}^{k2}| \qquad (۷۵)$$

y_i^{k2} و x_{ji}^{k2} برای $j=3$ رابر x_{2i}^{k2} برای تمام $i=1,\ldots,n$ تقسیم می کنیم،

$$y_i^{s2} = y_i^{k2}/x_{2i}^{k2} \qquad i=1,\ldots,n \qquad (۷۶)$$

$$x_{ji}^{s2} = x_{ji}^{k2}/x_{2i}^{k2} \qquad i=1,\ldots,n;\ j=3 \qquad (۷۷)$$

تابع ابژکتیو به شکل زیر می شود ،

$$S_{k1k2} = \sum_{i=1}^{n} |x_{1i}^{k1}x_{2i}^{k2}|\,|y_i^{s2}-\beta_2-\beta_3 x_{3i}^{s2}| \qquad (۷۸)$$

y_i^{s2} و x_{ji}^{s2} راازمشاهده $k3$ ام منحرف می کنیم ،

$$y_i^{k3} = y_i^{s2} - y_{k3}^{s2} \qquad i=1,\ldots,n \qquad (۷۹)$$

$$x_{ji}^{k3} = x_{ji}^{s2} - x_{jk3}^{s2} \qquad i=1,\ldots,n,\ j=3 \qquad (۸۰)$$

تابع ابژکتیو را به شکل زیربازنویسی میکنیم ،

$$S_{k1k2} = \sum_{i=1}^{n} |x_{1i}^{k1}x_{2i}^{k2}x_{3i}^{k3}|\,|y_i^{s3}-\beta_3 x_{3i}^{k3}| \qquad (۸۱)$$

y_i^{k3} رابر x_{3i}^{k3} برای تمام $i=1,\ldots,n$ تقسیم می کنیم

$$y_i^{s3} = y_i^{k3}/x_{3i}^{k3} \qquad i=1,\ldots,n \qquad (۸۲)$$

نهایتا " تابع ابژکتیو به شکل زیر درمی آید.

$$S_{k1k2k3} = \sum_{i=1}^{n} |x_{1i}^{k1}x_{2i}^{k2}x_{3i}^{k3}|\,|y_i^{s3}-\beta_3| \qquad (۸۳)$$

(۸۳) رابه شکل زیربازنویسی میکنیم

$$S_{k1k2k3} = \sum_{i=1}^{n} |w_i^{k1k2k3}|\,|r_i^{s1s2s3} - \beta_3| \qquad (۸۴)$$

کــه

$$w_i^{k1k2k3} = x_{1i}^{k1} x_{2i}^{k2} x_{3i}^{k3} \qquad (۸۵)$$

$$r_i^{s1s2s3} = y_i^{s3} \qquad (۸۶)$$

تابع ابژکتیو (۸۴) یك مسئله میانه وزنی است که براحتی برای پیداکردن β_3 میتواند حداقل شود٠معادله (۸۴) برای الگوی m پارامتری عمومی

$$y_i = \sum_{j=0}^{m-1} x_{ji}\beta_j + u_i \qquad i=1,\ldots,n \qquad (۷۸)$$

$$x_{0i} = 1 \qquad i=1,\ldots,n \qquad (۸۸)$$

باتابع ابژکتیو

$$S = \sum_{i=1}^{n} \left| y_i - \sum_{j=0}^{m-1} x_{ji}\beta_j \right| \qquad (۸۹)$$

برابراست بــا

$$S_{k1k2\ldots k(m-1)} = \sum_{i=1}^{n} |w_i^{k1k2\ldots k(m-1)}|\,|r_i^{s1s2\ldots s(m-1)} - \beta_{m-1}| \qquad (۹۰)$$

کــه

$$w_i^{k1k2\ldots k(m-1)} = |x_{1i}^{k1} x_{2i}^{k2} \ldots x_{(m-1)i}^{k(m-1)}| \qquad (۹۱)$$

$$r_i^{s1s2\ldots s(m-1)} = y_i^{s(m-1)} \qquad (۹۲)$$

حالا ،خواص الگوریتم ۴رابا استفاده ازفرمولاسیون فوق توضیح می دهیم

خاصیــــت ۱

اگرمحاسبات رابا $m-1$ نقطه که متعلق به مجموعه اصلی مشاهدات نیستندشــروع کنیم ،مقدار S درهردور به حداقل کلی خودش نزدیك می شود٠فرض کنیدنقاط

$(y_{n+1},\ldots,x_{m-1,n+1}),\ldots,(y_{n+m-1},\ldots,x_{m-1,n+m-1})$

به مجموعه نقاط نمونه ماتعلق ندارند٠اگر $n+1, n+2, \ldots, n+m-1$

به $k1, k2, \ldots, k(m-1)$ منصوب شوند،حداقل کردن تابع (۹۰) منجربه پیداشـــدن

نقطه M می شودکه‌متعلق به‌مشاهدات نمونه‌است ٠جایگزینی $k1$ با M وحذف نقطه ($n+1$) ام‌ازپایه‌وحداقل کردن(۹۰)منجربه‌پیدا شدن نقطه‌دیگری مانند M' متعلق‌به‌نمونه می شودکه‌جایگزین مشاهده $k2$ ام می شودکه خارج ازنمونه‌ما است ٠بدین ترتیب درهردوریکی ازنقاط خارج ازنمونه‌حذف وروش بهینه‌سازی بعداز $m-1$ دوربرروی نقاط نمونه‌انجام‌خواهد گرفت ٠

این خاصیت کاربردخوبی درداده‌های ill - conditioned دارد،به عبارت دیگر اگربرای مثال خطای گردکردن سبب این شودکه یك مشاهده مقدارنادرست گردشده به خود بگیرد، این نقطه دردوربعدبایك نقطه نمونه جایگزین خواهدشدوبدین ترتیب مسیرنزول رابه نقطه صحیح بازمی گرداند٠

خاصیت ۲

ساده کردن الگوی $m-1$ پارامتری به الگوئی باپارامترهای کمتررا میتوان با استفاده ازمحاسبات الگوی $m-1$ پارامتری وبلعکس انجام داد٠ بدین ترتیب رگرسیون های مرحله ای وتمام رگرسیون های ممکن را میتوان کاراترمحاسبه نمود٠

برای نشان دادن این خاصیت فرض کنیدبا معادله (۸۶) مواجه هستیم ٠ مقدار r_i^{s1s2s3} رامیتوان باجایگزینی (۶۷)،(۶۸)،(۷۰)،(۷۱)،(۷۳)،(۷۴)،(۷۶)،(۷۷)،(۷۹)،(۸۰)، (۸۱) ، (۸۲) در (۸۶) محاسبه کردنتیجه برابراست با ،

$$r_i^{s1s2s3} = \frac{\dfrac{\dfrac{y_i-y_{k1}}{x_{1i}-x_{1k1}} - \dfrac{y_{k2}-y_{k1}}{x_{1k2}-x_{1k1}}}{\dfrac{x_{2i}-x_{2k1}}{x_{1i}-x_{1k1}} - \dfrac{x_{2k2}-x_{2k1}}{x_{1k2}-x_{1k1}}} - \dfrac{\dfrac{y_{k3}-y_{k1}}{x_{1k3}-x_{1k1}} - \dfrac{y_{k2}-y_{k1}}{x_{1k2}-x_{1k1}}}{\dfrac{x_{2k3}-x_{2k1}}{x_{1k3}-x_{1k1}} - \dfrac{x_{2k2}-x_{2k1}}{x_{1k2}-x_{1k1}}}}{\dfrac{\dfrac{x_{3i}-x_{3k1}}{x_{1i}-x_{1k1}} - \dfrac{x_{3k2}-x_{3k1}}{x_{1k2}-x_{1k1}}}{\dfrac{x_{2i}-x_{2k1}}{x_{1i}-x_{1k1}} - \dfrac{x_{2k2}-x_{2k1}}{x_{1k2}-x_{1k1}}} - \dfrac{\dfrac{x_{3k3}-x_{3k1}}{x_{1k3}-x_{1k1}} - \dfrac{x_{3k2}-x_{3k1}}{x_{1k2}-x_{1k1}}}{\dfrac{x_{2k3}-x_{2k1}}{x_{1k3}-x_{1k1}} - \dfrac{x_{2k2}-x_{2k1}}{x_{1k2}-x_{1k1}}}} \qquad (۹۳)$$

بابررسی (۹۳) ،واضح است که محتویات خطوط مقطع‌گوشه‌دار، r_i^{s1} ، r_i^{s1s2} و r_i^{s1s2s3} هستند ٠بطورکلی $r_i^{s1s2...s(m-1)}$ شمای زیررادارا است ،

$$r_i^{s1\ldots s(m-1)} = \left.\underline{\left.\underline{\left.\underline{\left.\underline{r_i^{s1}}\right|\quad r_i^{s1s2}}\right|\quad \ddots}\right|\quad r_i^{s1\ldots s(m-1)}}\right| \tag{۹۴}$$

بحثی مشابه رانیزمیتوان برای $w_i^{k1k2\ldots k(m-1)}$ مطرح کرد. بـــــا جایگزینی (۶۸) و(۷۴) و (۸۰) در (۹۱) آنرابازنویسی می کنیم

$$w_i^{k1k2k3} = \left| (x_{1i}-x_{1k1})\left(\frac{x_{2i}-x_{2k1}}{x_{1i}-x_{1k1}} - \frac{x_{2k2}-x_{2k1}}{x_{1k2}-x_{1k1}}\right) \left(\frac{\dfrac{x_{3i}-x_{3k1}}{x_{1i}-x_{1k1}} - \dfrac{x_{3k2}-x_{3k1}}{x_{1k2}-x_{1k1}}}{\dfrac{x_{2i}-x_{2k1}}{x_{1i}-x_{1k1}} - \dfrac{x_{2k2}-x_{2k1}}{x_{1k2}-x_{1k1}}} - \frac{\dfrac{x_{3k3}-x_{3k1}}{x_{1k3}-x_{1k1}} - \dfrac{x_{3k2}-x_{3k1}}{x_{1k2}-x_{1k1}}}{\dfrac{x_{2k3}-x_{2k1}}{x_{1k3}-x_{1k1}} - \dfrac{x_{2k2}-x_{2k1}}{x_{1k2}-x_{1k1}}}\right)\right| \tag{۹۵}$$

بابررسی (۹۳) ،(۹۴) و(۹۵) واضح است که محاسبه $r_i^{s1s2\ldots sh}$ و $w_i^{s1s2\ldots sh}$ و $0<h\leq m-1$ کهبرای رگرسیون h پارامتـــری‌لازم هستندرامیتوان ازقدم های محاسباتی الگوی m پارامتری بدست آورد.

خاصیـــت ۳

بازاء هرجابجائی برای kj و $j=1,\ldots,m-1$ مقادیر $r_i^{s1s2\ldots s(m-1)}$ تعریف شده در(۹۲) برای تمام $i\neq k1,k2,\ldots,k(m-1)$ تـغییرنمی کنند. ایـــــن محاسبات باحذف جملات مناسب در(۹۳) و(۹۴) بعدازمخرج مشترك گرفتن می توانددانجــام شود.

بواسطه این خاصیت ،میتوانیم مقادیر $k1,\ldots,k(m-1)$ رادرفرآینـــد محاسبات بدون تغییرمقدار $r_i^{s1\ldots s(m-1)}$ جابجاکنیم.

خاصیـــت ۴

باهرگونه جابجائی kj برای $j=1,\ldots,m-1$ ،مقادیر $w_i^{k1\ldots k(m-1)}$ برای $i=1,\ldots,n$ در(۹۱) تغییر نخواهدکرد.

این نتیجه‌گیری ازضرب پرانتزهای متوالی(۹۷) وحذف جملات مناسب بدست خواهـد آمد. این خاصیت همراه باخاصیت ۳ تضمین می نمایدکه هرجابجائی $k1,\ldots,k(m-1)$ اقلام لازم(۹۱) و(۹۲) رادرمحاسبه مسئله میانه وزنـــی (۹۰) تغییرنخواهد داد.

خاصیت ۵

مقدار $r_{k(m-1)}s1...s(m-1)=0$ ، باجایگزینی $i=k(m-1)$ در (۹۳) یا (۹۴) این خاصیت نتیجه‌می شود. ازآنجائیکه این جمله مساوی صفراست بایدازچك تقسیم (divide check) (تقسیم برصفر) درفرآیندمحاسبات $r_{k1}s1, r_{k2}s1s2, ..., r_{k(m-1)}s1s2...s(m-1)$ آگاه بود.

خاصیت ۶

مقادیر $w_h k1k2...k(m-1)=0$ برای تمام $h=k1,k2,...,k(m-1)$ اگرمساوی یکی از $k1,k2,...,k(m-1)$ باشد، یکی ازپرانتزهای (۹۵) مساوی صفراست واین خاصیت بدست می آید. براساس این خاصیت برای اجتناب ازچك تقسیم، بدون محاسبه $w_h k1k2...k(m-1)$ برای $h=k1,k2,...,k(m-1)$ می توان مقدارآنها رامساوی صفرقرارداد.

خاصیت ۷

برای هرجابجائی kj برای $j=1,...,(m-1)$ ،نقطه M ام حاصل ازحداقل کردن (۸۹) ثابت باقی می ماند.زیرابااین جابجائی مقادیر $r_i s1s2...s(m-1)$ و $w_i k1k2...k(m-1)$ تغییرنخواهندکرد(خواص ۳، ۴، ۵، ۶).

خاصیت ۸

روشهای بهنگام کردن برای ارزیابی زیرمجموعه‌های مشاهدات نمونه‌براحتی میتوانند باصفرکردن مقادیر $w_h k1k2...k(m-1)$ آنهادر(۹۰) برای نقاطی که بایدحذف شوندانجام گیرد.این امرسبب می‌شودکه نقاط ناخواسته h درمحاسبات بی‌اثر باشند.

خاصیت ۹

برطبق خواص ۶ و ۸ ،ازآنجائیکه مقدار $w_h k1k2...k(m-1)=0$ برای تمام

، نقطه M حاصل ازحل مسئله میانه وزنی (۹۰) مساوی $h=k1,k2,...,k(m-1)$ نخواهدبود. این خاصیت تضمین می نمایدکه زمانیکه یک مشاهده $k1,k2,...,k(m-1)$ راواردپایه می کنیم این نقطه رامجددا" درهمان دورپیدانخواهیم کرد.

خاصیت ۱۰

اگرتمام kj هابرای $j=1,...,m-1$ درزمان شروع مساوی باشند، مسئله‌ای در فرآیندمحاسبات اتفاق نخواهدافتاد، ولی زمان اجراافزایش خواهدیافت. بلعکس اگر نقاط kj طوری انتخاب شوندکه نزدیک حداقل S باشند، الگوریتم بادورهای کمتری به‌جواب بهینه خواهدرسید.

خاصیت ۱۱

مقادیر $w_i k1k2...k(m-1)$ و $r_i s1s2...s(m-1)$ که برای حل مسئله میانه وزنی (۹۰) اساسی هستندرامیتوان ازمقادیرآنان دردورقبل محاسبه نمود. بدین معنی که‌زمانی که نقطه kh ازپایه بیرون می رودو kj جایگزین آن می شود، مقادیر $w_i k1k2...k(h-1)$ و $r_i s1s2...s(h-1)$ تغییرنمی کنندواحتیاج به محاسبه مجددآنهانیست ،بکارگیری این خاصیت درالگوریتم ۴ آنراخیلی کارآترمی‌نمایند.

خاصیت۱۲

ازآنجائیکه مقادیر $r_i s1s2...s(m-1)$ و $w_i k1k2...k(m-1)$ درهر دور بایدازداده‌های ورودی اصلی y و X محاسبه شوند، خطای گردکردن انباشته‌نمی شود وهمگرائی پریشان نمی گردد. این خاصیت سبب ایمن شدن الگوریتم ازانباشت خطای گرد شدن که درالگوریتم های نوع سیمپلکس وجوددارد‌میگردد.

خاصیت ۱۳

زمانی که m نقطه‌بهینه پیداشدند، مقادیربهینه $\hat{\beta}_j$ برای $j=1,...,m$ رامی توان ازیک دستگاه بازگشتی معادلات بطورکارامحاسبه نمود. مقادیر $\hat{\beta}_j$ برابر

خواهدبودبا ،

$$
\begin{bmatrix} \hat{\beta}_{m-1} \\ \hat{\beta}_{m-2} \\ \vdots \\ \hat{\beta}_2 \\ \hat{\beta}_1 \\ \hat{\beta}_0 \end{bmatrix}
=
\begin{bmatrix} y_{M}^{s(m-1)} \\ y_{k(m-1)}^{s(m-2)} \\ \vdots \\ y_{k3}^{s2} \\ y_{k2}^{s1} \\ y_{k1} \end{bmatrix}
-
\tag{96}
$$

$$
\begin{bmatrix}
0 & 0 & \cdots & 0 & 0 & 0 \\
x_{(m-1)k(m-1)}^{s(m-2)} & 0 & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & & \vdots & \vdots & \vdots \\
x_{(m-1)k3}^{s2} & x_{(m-2)k3}^{s2} & \cdots & 0 & 0 & 0 \\
x_{(m-1)k2}^{s1} & x_{(m-2)k3}^{s1} & \cdots & x_{2k2}^{s1} & 0 & 0 \\
x_{(m-1)k1} & x_{(m-2)k1} & \cdots & x_{2k1} & x_{1k1} & 0
\end{bmatrix}
\begin{bmatrix} \hat{\beta}_{m-1} \\ \hat{\beta}_{m-2} \\ \vdots \\ \hat{\beta}_2 \\ \hat{\beta}_1 \\ 0 \end{bmatrix}
$$

دستگاه بازگشتی معادلات (۹۶) بشکل زیربدست می آیند. (۶۷) و (۶۸) رابرای y_i و x_{ji} ، $j=1,2,3$ حل کرده ودر (۶۶) جایگزین می کنیم

$$S_{k1} = \sum_{i=1}^{n} \left| y_i^{k1} - \beta_1 x_{1i}^{k1} - \beta_2 x_{2i}^{k1} - \beta_3 x_{3i}^{k1} - (y_{k1} - \beta_0 - \beta_1 x_{2k1} - \beta_3 x_{3k1}) \right| \tag{97}$$

ازآنجائیکه رگرسیون راملزم به عبورازنقطه $k1$ کرده ایم ،عبارت داخل پرانتـــز (۹۷) مساوی صفراست .بنابراین

$$y_{k1} - \beta_0 - \beta_1 x_{2k1} - \beta_3 x_{3k1} = 0 \tag{98}$$

می توانیم (۷۳) و (۷۴) رادوباره برای y_i^{s1} و x_{ji}^{s1} که $j=2,3$ حل کنیـم و آنهارادر (۷۲) جایگزین نمائیم کەمنجربەتابع ابژکتیوزیرمی شود.

$$S_{k1k2} = \sum_{i=1}^{n} \left| x_{1i}^{k1} \right| \left| y_i^{k2} - \beta_2 x_{2i}^{k2} - \beta_3 x_{3i}^{k2} - (y_{k2}^{s1} - \beta_1 - \beta_2 x_{2k2}^{s1} - \beta_3 x_{3k2}^{s1}) \right| \tag{99}$$

بابحثـی مشابەدر (۹۷) داریم ،

$$y_{k2}^{s1} - \beta_1 - \beta_2 x_{2k2}^{s1} - \beta_3 x_{3k2}^{s1} = 0 \tag{100}$$

همینطورمی توانیم بدست آوریم

$$S_{k1k2k3} = \sum_{i=1}^{n} \left| x_{1i}^{k1} x_{2i}^{k2} \right| \left| y_i^{k3} - \beta_3 x_{3i}^{k3} + (y_{k3}^{s2} - \beta_2 - \beta_3 x_{3k3}^{s2}) \right| \tag{101}$$

وبرمبنای (۱۰۱)

$$y_{k3}^{s2} - \beta_2 - \beta_3 x_{3k3}^{s2} = 0 \tag{102}$$

نهایتا " ،باحل مسئله میانه وزنی (۸۲) نقطه M رامی یابیم که میدهد

$$\beta_3 = y_{M}s3 \qquad (۱۰۳)$$

معادلات (۱۰۳) ،(۱۰۲) ،(۱۰۰) و(۹۸) رامیتوان پشت سرهم برای $\hat{\beta}_0, \ldots, \hat{\beta}_3$ حل نمود.دریك شكل مشابه جواب عمومی برای الگوی m پارامتری در(۹۶) ارائه‌گردیده است .حالااین خاصیت، هارامورداستفاده قرارداده‌ومراحل محاسباتی الگوریتم ۴ راخیلی مشروح تردربرنامه BL 1 زیرمی بینیم.این‌برنامه‌درفصل بعددرمقایسه‌باسایرالگوریتمها برای رگرسیون‌نرم L_1 عمومی مورداستفاده قرارخواهدگرفت

```
PROGRAM BL1
Step 0) Initialization.
        Parameter: n, m, m1=m-1, m2=m-2.
        Real: y(n), x(n,m1), xsk(m1), yw(n), xkw(n), w(n),
        ys(n), xs(n,m1), b(m), xw(n,2:m1), ysol(m1),
        xsol(m1,m1).
        Integer: l(n), kk(m1).
        Common: /c1/i1,i2.
        Read: (y(i),(x(i,j),j=1,m1),i=1,n).
        Let: iter=0, kr=0, mm=1, (kk(j)=arbitrary,j=1,m1).
Step 1) Refill working arrays.
        Do loop for i=1,n: ys(i)=y(i), do loop for j=1,m1:
        xs(i,j)=x(i,j), end do, end do.
Step 2) Store weights and ratios for next iteration.
        Do loop for i=1,n: w(i)=xkw(i), ys(i)=yw(i),  do loop
        for j=1,m1: xs(i,j)=xw(i,j), end do, end do.
Step 3) Compute the arguments for weighted median.
     a. Set: jj=mm, k=kk(jj), ysk=ys(k), i1=1, i2=k-1.
        Do loop for j=jj,m1: xsk(j)=xs(k,j), end do.
     b. Do loop for j=jj,m1: call COL1(xsk(j),xs(1,j)) end
        do.
        Call COL2(ysk,jj,w,ys,xs(1,jj)).
        If i2=n go to step 3.c; otherwise set:  i1=k+1, i2=n,
        go to step 3.b.
```

c. Set: w(k)=0.

If jj=m1 go to step 4; otherwise i1=1, i2=k-1, go to step 3.d.

d. Do loop for j=jj+1,m1: call COL3(xs(1,j),xs(1,jj)), end do.

If i=n go to step 3.e; otherwise i1=k+1, i2=n, go to step 3.d.

e. If jj≠mm jj=jj+1, go to step 3, otherwise do loop for i=1,n: xkw(i)=w(i), yw(i)=ys(i); do loop for j=jj+1,m1: xw(i,j)=xs(i,j), end do; end do.

Set: jj=jj+1, go to step 3.

Step 4) Compute the weighted median.

Set: ys(k)=0, iter=iter+1, lm=LWMED(n,ys,w,1).

Step 5) Test for optimality.

If lm=kr go to step 5.b; otherwise iopt=0 go to step 5.a.

a. If mm=m1 set mm=1, kr=kk(mm), kk(mm)=lm, go to step 1; otherwise set mm=mm+1, kr=kk(mm), kk(mm)=lm, go to step 2.

b. Set: iopt=iopt+1.

If iopt≠m1 go to step 5.a, otherwise go to step 6.

Step 6) Compute the solution.

Set: b(m)=ys(lm).

Do loop for i=1,m1: ysol(i)=y(kk(i)); do loop for j=1,m1: xsol(i,j)=x(kk(i),j), end do; end do.

Set: jj=1.

a. Set: ysk=ysol(jj).

Do loop for j=jj,m1: xsx(j)=xsol(jj,j), end do.

Do loop for i=jj,m1: if i=jj go to continue; otherwise ysol(i)=ysol(i)-ysk, do loop for j=jj,m1: xsol(i,j)=xsol(i,j)-xsk(j), end do; set ysol(i)=ysol(i)/xsol(i,jj), continue, end do.

b. Do loop for i=jj,m1: if i=jj go to continue, otherwise, do loop for j=jj+1,m1: xsol(i,j)=xsol(i,j)/xsol(i,jj), end do; continue; end do.

```
c. If jj=m2 go to step 6.d; otherwise go to step 6.a.
d. Do loop for i=1,m2: k=m-i,  s=ysol(k);   do loop  for
   j=k,m1, s=s-b(j+1)*xsol(k,j) end do, b(k)=s, end do.
   Set: s=y(kk(1)).
   Do loop for j=1,m1:  s=s-b(j+1)*x(kk(1),j),   b(1)=s,
   end do.
   Print: ((b(j),j=1,m),(kk(j),j=1,m1),lm,iter).
   Stop.
END
```

سهم وسیعی ازمحاسبات دراین برنامه‌تبدیل آرایه‌های دوبعدی است ۰انتقال ستونهای این آرایه‌هابه‌سابرروتین های دیگرکه‌فقط آرایه‌های یك بعدی دارندزمان محاسبه راکاهش می دهد(نگاه کنیدبه Roberts ,Barrodale (1974)۰ سابروتین‌های COL1 و COL2 ,COL3 بدین منظوربترتیب برای تفریق ،ضرب وتقسیم ،وبرای فقط تقسیم نوشته شده‌اند۰تابع LWMED که‌برای محاسبه‌میانه وزنی بکاربرده می شوددرقسمت ۱۰۲ معرفی شده‌است ۰

```
SUBROUTINE COL1(v1,v2)
Step 0) Initialization
        Real: v2(1).
        Common /c1/i1,i2.
Step 1) Subtraction.
        Do loop for i=i1,i2: v2(i)=v2(i)-v1, end do.
        Return.
END

SUBROUTINE COL2(ysk,jj,v1,ys,v2)
Step 0) Initialization.
        Real: v1(1),v2(1),ys(1).
        Common /c1/i1,i2.
Step 1) Compute weights and ratios.
        If jj≠1 go to step 1.a,; otherwise do loop for
        i=i1,i2: v1(i)=v2(i), ys(i)=(ys(i)-ysk)/v2(i).
        Return.
     a. Do loop for i=i1,i2: v1(i)=v1(i)*v2(i),
        ys(i)=(ys(i)-ysk)/v2(i), end do.
        Return.
END
```

```
SUBROUTINE COL3(v1,v2)
Step 0) Initialization.
        Real: v1(1),v2(1),ys(1).
        Common /c1/i1,i2.
Step 1) Division.
        Do loop for i=i1,i2: v1(i)=v1(i)/v2(i), end do.
        Return.
END
```

۴-۴- مسئله مقدار اولیه

انتخاب نقاط دلخواه k1,k2,.....,k(m-1) در ابتدای محاسبات نقش مهمی در تعداد دور تکرار لازم برای رسیدن به جواب بهینه دارد. یک راه برای انتخاب این m - 1 نقطه می‌تواند بر مبنای بکارگیری برآوردکننده دیگری باشد تا حدس بزند کدام نقاط دارای قدر مطلق پسماند کوچکترین در میان تمام پسماندهای برآورد شده هستند. بدین معنی که یک برآوردکننده مثلاً " حداقل مربعات را در مورد داده‌ها بکار می‌بندیم و m - 1 پسماند را با کوچکترین قدر مطلق انتخاب می‌کنیم و نقاط مربوط به این پسماندها را برای شروع الگوریتم‌های ۱، ۲، ۳ یا ۴ نامزد می‌نمائیم. به هرحال، این روش برای الگوریتم‌های ۳ و ۴ مورد آزمایش قرار گرفت و سبب کاراتر شدن آنان گردید. جزئیات آزمایشات درج نمی‌شود.

فصل چهارم

" مقایسه الگوریتم ها "

۱. مقـــــدمـــــه

هدف این فصل مقایسه بعضی ازالگوریتم هائی که درفصل دوم توضیح داده شدبـــــا الگوریتم های پیشنهادی درفصل سوم می باشد. نقطه نظرما مقایسه صحت وکارآئی نسبــی آنهاست . ازاین جهت صحت جواب الگوریتم ها مهم ترازسایرمعیارهاست .ازواژه صحـــت منظورمارسیدن به جواب درست درتعدادی محدودازدورتکراریامرحله است .منظورازکارآئی این است که الگوریتم موردنظربا میزان کمتری ازفضای حافظه وزمان اجرابه جواب بهینـــه درست دست یابد.

درقسمت ۲ طرح آزمایشات موردبحث است . تولیددادههای لازم ،تشخیص الگوهـــابـــا تعدادپارامترهاومشاهدات مربوطه ومحیط محاسباتی موضوعات این قسمت می باشنـــــد. قسمت ۳ الگوریتم های نرم L_1 رابرای رگرسیون ساده خطی مقایسه مینماید.درایـــن حالت الگوریتم ۲ فصل قبل باالگوریتم (1983) Sposito و Joswanger مقایسه می شوند. درقسمت ۴ ،الگوریتم های نرم L_1 برای رگرسیون چندمتغیره خطــی ملاحظه خواهندشد. دراین جهت ،الگوریتم ۴ فصل سوم باالگوریتم های (1973) Armstrong , Frome , Kung (1979) Barrodale , Roberts (1980) Bloomfield,Steiger کهازسایرالگوریتم های مذکوردرفصل دوم صحیح تر وکارآترمی باشندمقایسه می شوند.قسمت آخربه نتیجهگیری اختصاص خواهدداشت .

۲. طـــرح آزمایشـــــات

عملکردهرالگوریتمی درهرمحیط محاسباتی خاص متفاوت است وبنابراین مقایسهمطلق الگوریتم هاخیلی مشکل می باشد ،مخصوصا " اگرسیستم موردنظرازحافظه حقیقی یامجـــازی ،ازیک Cache یاهرآرایه پردازنده دیگر ،کمک پردازنده ریاضی ویاغیره استفاده نمایــد. همانطورکه درفصل دوم بحث شد ،الگوریتم های زیادی برای رگرسیون نرم L_1 همراه بـــا برنامههای کامپیوتری مربوطه وجوداردومقایسه تمام آنهاکاری پرهزینه میباشد. بـــرای اینکه تعدادآزمایشات راکاهش دهیم برتجربه سایرمحققین که درفصل دوم موردبحث قـــرار گرفت تکیه می کنیم . به هرحال ،آزمایشات به دوگروه عمده رگرسیون های نرم L_1 خطــی ساده وچندمتغیره تقسیم می شوند.

علیرغم برنامه های کامپیوتری نوشته شده ،محیط محاسباتی ،تعدادمشاهدات و پارامترهای الگوو " شرایط " داده هامنبع اصلی مقایسه عملکردالگوریتم هامی باشد.بدین ترتیب مسئله‌هائی با اندازه‌های مختلف دراین قسمت موردآزمایش قرارمی گیرند.

برای قضاوت درموردبرتری الگوریتم ها معیارهای مختلفی وجوددارد.صحت وکارآئی از مهمترین آنان می باشند. دراولی منظوربدست آوردن نتایج درست درنمونه‌های مختلف و دردومی زمان وفضای حافظه لازم برای محاسبه الگوریتم های موردمقایسه مدنظرمی باشد.

برای انجام آزمایشات ،یکبارمقادیرتصادفی یکنواخت برای β_j درالگوی زیر انتخاب شوند،

$$y_i = \sum_{j=1}^{m} \beta_j x_{ij} + u_i \qquad i=1,...,n \tag{1}$$

مقادیرتصادفی برای x_{ij} و u_i باپنج تشخیص توزیع تولیدشد. تولیدکننده‌های تصادفی یکنواخت ونرمال (داده شده توسط (1977 Mojarrad) برای تولیدسه مجموعه ازداده‌های تصادفی یکنواخت ودومجموعه ازداده‌های تصادفی نرمال برای هر آزمایش مورداستفاده قرارگرفت .اعدادتصادفی یکنواخت به فواصل $[-10,10]$ ، $[-100,100]$, $[-1000,1000]$ تعلق دارند.اعدادنرمال بامیانگین صفر وواریانس 1000,100 انتخاب شدند.مقادیر y_i برای x_{ij} , β_j و u_i که همانطورکه دربالاگفته شدتولیدشده بودندمحاسبه گردید.مقادیر ۲۰، ۵۰، ۱۰۰، ۵۰۰، ۱۰۰۰، ۲۰۰۰ ، ۵۰۰۰و ۱۰۰۰۰ برای تعدادمشاهدات n استفاده شدندومقادیر ۲، ۳، ۴، ۵، ۷، ۱۰ رابرای تعدادپارامترها m انتخاب گردیدند.

بدین ترتیب ،برای تمام پنج تشخیص توزیع u_i و برای تمام m ها , n ها یک آزمایش برای هرکدام ازالگوریتم ها انجام می شود. میانگین ودامنه این پنج آزمایش برای هر m و n وبرای هرالگوریتم گزارش می شوند.درحالت رگرسیون ساده تعدادآزمایشات بجای پنج ده می باشد.

تمام برنامه‌هابافورترن IV ، کمپایلر VS (level 1.3.0 (۱۹۸۳) و LANGLVL 77 وسطح بهینه سازی 03 برای کاهش ناکارآئی برنامه‌های نوشته شده کمپایل گردیدند.برنامه‌هاهمگی برروی کامپیوتر 7.68 BASF (MVS) اجراشدند. ازآنجائیکه این ماشین یك سیستم چندکاره است ،فرآیندهای بیرون وداخل کردن

(Swapping) برزمان اجراتاثیرمی گذارند ٠زمانیکه سیستم برای بیش ازیك جاب (job) درحال اجراست این اثرراکاملا" نمی توان اندازه گیری وازبین برد٠ برای تصفیه زمان بیرون وداخل کردن ، زمان بلوك درخواست سیستم (SRB) ازکل زمان واحد پردازش‌مرکزی (CPU) کسرگردیده‌است ٠ به هرحال ز مانیکه سیستم مشغول است این کارکاملا" تمام زمانهای بیرون وداخل کردن راتصفیه نمی نماید ٠سعی گردیده‌که کلیه الگوریتم های قابل مقایسه دریك زمان ودریك کلاس ورودی باinitiator های‌کافی وسطح اولویت یکسان برای ایجادوضعیت مشابه برای‌تمام جاب های واردشده قابل مقایسه اجرا گردند ٠زمان پیش ازاجرای کمپایل و edit - Link ازتمام برنامه‌های موردآزمون حذف شده اند٠

٠٣ مقایسه الگوریتم‌های نرم L_1 رگرسیون ساده

دراین مطالعه مقایسات محدودبه الگوریتم ۲ ازفصل قبل والگوریتم پیشنهادی (1983)Josvanger,Sposito محدودمی شوند٠ (1987) Sposito , Gentle , Narula (1988) Gentle , Sposito , Narula این الگوریتم راکاراترین الگوریتم برای رگرسیون نرم L_1 خطی ساده معرفی می نمایند٠ مطالعات آنان الگوریتم های (1958) Karst بابرنامه (1974) Sadovski ،(1978)Kung Armstrong (1979) Kung , Frome Armstrong (1980) Bloomfield ,Steiger (1981) Wesolowsky (1983) Sposito , Josvanger رامقایسه می نماید٠

جدول ۱ : نیازفضای آرایه برای الگوریتم های منتخب الگوی ساده

Algorithm	Program name	Storage requirement	Stopping constant(s)
JS	DESL1	5n	TOL = 1.0E-6
B(Alg.2)	BL1S	5n	--------------

Js ≡ Josvanger and Sposito (1983).
B(Alg.2) ≡ Bidabad (Algorithm 2) chapter III.
n ≡ Number of observations.

میزان نیازحافظه آرایه برای این دوبرنامه در جدول ۱ آورده شده است. این جدول رامیتوان باجدول ۲ درفصل دوم برای سایرالگوریتم ها مقایسه نمود. هیچکدام ازبرنامه ها داده های ورودی راخراب نمی کنند. هردوبرنامه بادقت مفردنوشته شده اند.

جدول ۲ نتایج آزمایشات برای رگرسیون نرم L_1 خطی ساده رانشان می دهد. مقادیر گزارش شده درسلولهای جدول میانگین های زمانهای CPU ده آزمایش درنمونه های تصادفی مختلف به ثانیه می باشند. مقادیردرون پرانتزهازمانهای CPU حداقل وحداکثر مربوطه ده آزمایش می باشند. هردوالگوریتم درتمام آزمایشات به جوابهای درست دست یافتند.

همانطورکه ازجدول ۲ مشهوداست درنمونه های کوچك گرچه الگوریتم ۲ سریعتر میباشد زمانهای محاسبه خیلی متفاوت نیستند. درنمونه های متوسط، این تفاوت با معنی می شودودرنمونه های بزرگترالگوریتم ۲ قویا " برترازالگوریتم (Sposito 1983) Josvanger می شود. بنابراین میتوان نتیجه گرفت که الگوریتم ۲ بهتر از سایرالگوریتم ها عمل می نمایدومیتوان ازآن درکارهای عملی برای حصول کارآئی بیشتر استفاده نمود.

جدول ۲ زمانهای CPU برای الگوریتمهای منتخب نرم L_1 برای الگوی ساده

n	JS	B(Alg.2)
20	0.096 (0.09,0.10)	0.094 (0.09,0.10)
50	0.109 (0.10,0.11)	0.106 (0.10,0.11)
100	0.141 (0.13,0.16)	0.132 (0.12,0.14)
500	0.469 (0.35,0.59)	0.360 (0.33,0.38)
1000	0.997 (0.66,1.31)	0.645 (0.61,0.69)
2000	2.770 (1.36,4.27)	1.194 (0.98,1.28)
5000	11.554 (4.58,18.91)	2.848 (2.71,3.15)
10000	42.406 (9.67,60.88)	5.823 (5.16,6.87)

نگاه کنید به جدول ۱ برای اختصارات.

۴. مقایسه الگوریتم های نرم L_1 رگرسیون چندمتغیره

برای مقایسه الگوریتم ۴ فصل سوم با سایر الگوریتم ها، آزمایشات درسه الگوریتمی که درمیان سایرین صحیح تروسریع ترمی باشند محدود شده است. این سه الگوریتم عبارتند از الگوریتم های (Barrodale , Roberts (1973 , 74) (BR) و (1980) Bloomfield , Steiger (BS) ، (1979) Armstrong , Frome , Kung (AFK) . اگرچه ، الگوریتم های BS و AFK سریع تراز BR هستند ، دلیل انتخاب الگوریتم BR این بودکه دوالگوریتم دیگرموفق به تولیدجواب صحیح برای نمونه های بزرگترنشدند(نگاه کنیدبه ، (1987) Gentle, Narula , Sposito وهمچنین به فصل دوم).

میزان نیازفضای حافظه آرایه برای این برنامه‌ها درجدول ۳ درج گردیده است. این جدول رامیتوان باجدول ۲ فصل دوم برای سایرالگوریتم ها مقایسه نمود. تمام برنامه‌ها بادقت مفردنوشته شده اند. هیچکدام ازبرنامه‌ها داده‌های ورودی راخراب نمی کنند.

جدول ۳: نیازفضای حافظه آرایه برای الگوریتم های منتخب الگوی چندمتغیره

Algorithm	Program name	Storage requirement	Stopping constant(s)
AFK	AFKL1	6n+m(n+3m/2+15/2)	ACU = 1.0E-6 BIG = 1.0E+15
BR	L1BAR	3n+m(n+5)+4	BIG = 1.0E+75
BS	BLOD1	4n+m(2n+4)	--------------
B(Alg.4)	BL1	2n+m(3n+m+2)-2	--------------

AFK ≡ Armstrong and Frome and Kung (1970).
BR ≡ Barrodale and Roberts (1973,74).
BS ≡ Bloomfield and Steiger (1980).
B(Alg.4) ≡ Bidabad (Algorithm 4) chapter III.
n ≡ Number of observations.
m ≡ Number of parameters.

جداول ۴ تا ۸ میانگین های زمان CPU پنج اجرابرای اندازه های مختلف، نمونه‌ها وپارامترهاراگزارش می نماید. مقادیرداخل پرانتزها حداقل وحداکثرزمان CPU آزمایشات می باشند. برای الگوی سه پارامتری ، همانطورکه ازجدول ۴ مشهوداست، الگوریتم ۴ برترازسایرالگوریتم هامی باشد. دراین حالت الگوریتم های BS ، AFK و BR

بترتیب ازکارآئی کمتری برخوردارمی باشند. زمانیکه اندازه نمونه کوچك است ،اختلاف زیادنیست .دراندازه‌های نمونه متوسط ،این اختلاف افزایش می یابد. در آزمایشات اندازه های بزرگتر ، الگوریتم های ۴ و BS اختلاف کمی دارند،ولی BR و AFK خیلی ازآنهادورهستند.درتمام حالات الگوریتم ۴ سریعتر ازسایرالگوریتم هامیباشد.

جدول ۴: زمان CPU برای الگوریتم های منتخب الگوی چندمتغیره(m= 3)

n	B(Alg.4)	BR	BS	AFK
20	0.098 (0.09,0.10)	0.110 (0.11,0.11)	0.104 (0.10,0.11)	0.112 (0.11,0.12)
50	0.144 (0.13,0.16)	0.148 (0.14,0.16)	0.146 (0.13,0.16)	0.146 (0.14,0.15)
100	0.182 (0.18,0.19)	0.216 (0.20,0.23)	0.194 (0.19,0.20)	0.214 (0.20,0.23)
500	0.698 (0.63,0.74)	1.116 (1.07,1.17)	0.810 (0.63,1.00)	0.986 (0.85,1.11)
1000	1.390 (1.27,1.53)	2.420 (2.16,2.76)	1.662 (1.37,2.01)	2.180 (2.01,2.36)
2000	2.812 (2.34,2.99)	5.884 (4.98,6.63)	2.932 (2.81,3.18)	4.800 (4.30,5.09)
5000	7.456 (6.82,9.03)	25.038 (22.33,27.54)	7.520 (6.16,10.12)	20.172 (16.57, 22.33)
10000	14.330 (12.45,16.61)	80.008 (73.16,87.03)	15.434 (12.88, 18.12)	59.634 (55.60, 65.80)

نگاه کنیدبه‌جدول ۳ برای اختصارات .

درحالت الگوی چهارپارامتری همانطورکه درجدول ۵ نشان داده شده است ،گرچه الگوریتم BS رقیب الگوریتم ۴ می باشد، این رتبه بندی تغییرنمی کندوالگوریتم ۴ دوباره سریعترین می باشد.رتبه بندی الگوریتم های مختلف، شبیه آزمایشات سه پارامتری برای‌تمام حالات اندازه های نمونه کوچك ،متوسط وبزرگ می باشد.

زمانیکه‌تعدادپارامترهابه پنج افزایش یافت ،الگوریتم BS موفق به محاسبه جوابهای صحیح برای اندازه‌های نمونه ۲۰۰۰ وبیشترنگردید. (1987) Narula , Sposito و Gentle نیزبه شکست الگوریتم BS برای نمونه های ۱۰۰۰ وبیشتربرای الگوهای پنج پارامتری و بیشتروبرای نمونه‌های۵۰۰۰ بادوپارامتراشاره نمودند(همچنین نگاه کنید به فصل دوم).بارجوع به جدول (۶) باتوجه به شکست BS کارآئی الگوریتم ۴ نسبت به‌سایر

الگوریتم ها مشهوداست . الگوریتم های AFK و BK بترتیب دررتبه های بعدی قــرار دارند. برای نمونه های با اندازه کوچکتر ،الگوریتم های BR , BS , AFK

جدول ۵ زمان CPU برای الگوریتم های منتخب الگوی چندمتغیره (m=4)

n	B(Alg.4)	BR	BS	AFK
20	0.112 (0.11,0.12)	0.116 (0.11,0.12)	0.116 (0.11,0.12)	0.116 (0.11,0.12)
50	0.156 (0.15,0.16)	0.168 (0.16,0.19)	0.160 (0.15,0.17)	0.160 (0.16,0.16)
100	0.284 (0.26,0.32)	0.286 (0.26,0.30)	0.286 (0.27,0.30)	0.286 (0.26,0.30)
500	1.098 (0.85,1.44)	1.596 (1.23,1.77)	1.260 (0.92,1.58)	1.394 (1.17,1.71)
1000	2.194 (2.03,2.45)	4.016 (3.35,5.05)	2.200 (0.64,3.28)	3.022 (2.73,3.21)
2000	4.650 (3.92,5.05)	10.636 (9.44,11.86)	5.430 (4.54,6.15)	7.774 (7.25,8.16)
5000	12.852 (10.00,15.23)	41.282 (32.40,47.86)	12.938 (11.62, 14.04)	32.110 (29.55, 33.17)
10000	27.720 (22.93,39.14)	119.152 (107.14, 129.70)	27.864 (23.94, 32.10)	101.472 (100.31, 103.15)

برای اختصارات نگاه کنیدبه جدول ۳.

جدول ۶ زمان CPU برای الگوریتم های منتخب الگوی چندمتغیره (m=5)

n	B(Alg.4)	BR	BS	AFK
20	0.138 (0.13,0.15)	0.124 (0.12,0.13)	0.124 (0.12,0.14)	0.124 (0.11,0.13)
50	0.208 (0.18,0.24)	0.240 (0.22,0.26)	0.204 (0.20,0.21)	0.188 (0.18,0.20)
100	0.348 (0.33,0.37)	0.380 (0.34,0.42)	0.404 (0.36,0.46)	0.338 (0.32,0.36)
500	2.024 (1.57,2.40)	2.498 (2.29,2.68)	1.754 (1.34,2.01)	1.684 (1.57,1.78)
1000	3.702 (3.19,4.45)	5.684 (4.86,6.28)	3.364 (2.93,3.71)	3.876 (3.67,4.13)
2000	8.770 (7.51,9.41)	15.500 (13.04,16.58)	+ +	9.120 (7.91,9.93)
5000	24.418 (20.15,27.96)	66.394 (59.93,71.90)	+ +	36.600 (33.95, 38.67)
10000	53.924 (38.35,64.90)	244.072 (217.81, 270.06)	+ +	108.406 (99.65, 119.75)

+ موفق به محاسبه جواب صحیح نشده .

برای اختصارات نگاه کنیدبه جدول ۳.

جدول ۷ زمـــــان CPU برای الگوریتم های مختلف چندمتغیـــــره (m=7)

n	B(Alg.4)	BR	BS	AFK
20	0.160 (0.16,0.16)	0.164 (0.16,0.17)	0.156 (0.15,0.16)	0.160 (0.15,0.17)
50	0.346 (0.29,0.37)	0.302 (0.29.0.31)	0.328 (0.31,0.34)	0.282 (0.26,0.30)
100	0.706 (0.58,0.81)	0.572 (0.52,0.65)	0.540 (0.48,0.62)	0.530 (0.47,0.62)
500	3.898 (3.25,4.54)	4.486 (4.03,4.85)	3.202 (2.41,3.87)	2.990 (2.45,3.48)
1000	9.178 (7.03,10.44)	11.448 (9.88,12.71)	+ +	6.568 (5.87,7.68)
2000	19.984 (18.06,21.55)	31.872 (24.52,35.22)	+ +	14.908 (13.56, 16.23)
5000	57.286 (49.53,64.24)	141.234 (129.20, 154.03)	+ +	58.314 (49.12, 65.17)
10000	130.79 (103.20, 183.45)	475.826 (421.40, 521.39)	+ +	151.526 (144.94, 165.62)

+ موفق به محاسبه جواب صحیح نشده.

برای اختصارات نگاه کنیدبه جدول ۳.

جدول ۸ زمـــان CPU برای الگوریتم های منتخل الگوی چندمتغیو(m=10)

n	B(Alg.4)	BR	BS	AFK
20	0.276 (0.27,0.29)	0.212 (0.20,0.23)	0.218 (0.21,0.22)	0.212 (0.20,0.23)
50	0.956 (0.69,1.72)	0.502 (0.47,0.54)	0.492 (0.43,0.54)	0.398 (0.35,0.44)
100	2.414 (1.55,4.78)	1.144 (1.03,1.25)	0.998 (0.86,1.11)	0.776 (0.66,0.91)
500	13.980 (11.76,16.30)	8.446 (7.88,9.99)	5.970 (4.81,6.92)	5.210 (4.43,5.78)
1000	62.624 (22.23, 193.19)	23.506 (20.26,26.93)	+ +	11.144 (9.08,13.14)
2000	109.268 (72.20, 278.27)	62.756 (59.43,65.66)	+ +	+ +
5000	409.438 (154.64, 1010.92)	284.618 (240.02, 322.47)	+ +	+ +
10000	679.540 (283.04, 1076.04)	967.794 (770.79, 1064.43)	+ +	+ +

+ موفق به محاسبه جواب صحیح نشده.

برای اختصارات نگاه کنیدبه جدول ۳.

رقیب هستندولی اختلافات خیلی جزئی می باشند. درنمونه های بزرگتر، الگوریتم ۴ مشخصا" برترازسایرالگوریتم هامی باشد.

درجدول ۷ زمانیکه تعدادپارامترهاهفت می شود، الگوریتم BS موفق به محاسبه جواب صحیح برای نمونه های ۱۰۰۰ وبزرگترنشد. AFK برای نمونه های کوچکتر بهترین است، ولی برای نمونه های بزرگ، الگوریتم ۴ مجددا" برترازدیگران است. الگوریتم BR دررتبه سوم قرارمی گیرد.

درجدول ۸، باده پارامتر، الگوریتم های BS و AFK موفق به محاسبه جواب صحیح برای نمونه های بزرگترنشدند. الگوریتم BR باتوجه به کارآئی وصحت جواب بهترین است. الگوریتم ۴ بجزبرای نمونه ۱۰،۰۰۰ که سریعترین میباشدرتبه دوم رااز لحاظ زمان محاسبه وهم صحت به خوداختصاص می دهد.

۵. نتایج

ازآنجائیکه درالگوریتم های محاسباتی صحت مهم ترازکارآئی میباشد، آن الگوریتم های نرم L_1 رابایدانتخاب نمودکه قادربه محاسبه جواب صحیح هستندوازمیان آنها سریعترین آنهاراانتخاب کرد. الگوریتم ۲ والگوریتم (1983) Sposito و Josvanger هردوبرای الگوی رگرسیون نرم L_1 دوپارامتری جواب های صحیح محاسبه کردند. الگوریتم ۲ که سریعتراز JS می باشدبرای کارهای عملی معرفی می شود.

برای رگرسیون چندمتغیره الگوریتم های BS و AFK موفق به محاسبه جواب صحیح درالگوهای بزرگترنشدند. همانظورکه (1987) Gentle , Narula , Sposito ابرازمی نمایند، بدلیل انباشت خطای گردشدن، الگوریتم (1980) Steiger و Bloomfield درمسائل بزرگترقابل استفاده نبود. تغییربرنامه جهت اجتناب ازمسئله گردشدن اغلب زمان اجراراافزایش می دهد، بنابراین روشن نیست که باتغییربرنامه BS چه اتفاقی درموردکارآئی نسبی آن خواهدافتاد. این مسئله نیزدر موردالگوریتم (1979) Armstrong , Frome ,Kung صادق است اگرچه این الگوریتم

کمتراز BS به خطای گردشدن حساس است . به هرحال ازجدول قبل میتوان نتیجه گرفت
که الگوریتم ۴ برای الگوهای باکمتراز ده پارامتروالگوریتم (1973) Roberts
Barrodale برای الگوی ده پارامتری مناسب ترمی باشند. این نتیجه آخرخیلی
قطعی نمی باشد،زیردرحالت الگوی ده پارامتری با ۱۰٫۰۰۰ مشاهده الگوریتم ۴ خیلی برتر
از BR است . به هرحال ،ازآنجائیکه درکارهای عملی همیشه باتعدادبسیارزیادپارامترها
ومشاهدات روبرونیستیم ،این نتیجه گیری ازلحاظ عملی ضعیف است .

فصل پنجم

" برازش نرم L_1 پیوسته ومنحنی لورنز "

۱ـ مقدمه

درفصول قبل برآوردنرم L_1 داده‌های گسسته موردبحث قرارگرفت وموضوعات مرتبط باآن رابررسی کردیم . دراین فصل سعی براین است که بااستفاده ازچندتکنیکی که‌قبلا" توضیح داده شدمسئله برازش نرم L_1 توابع‌رابرای داده‌های پیوسته عملی نمائیم.کاربرد این گونه داده‌هادرمسائل برآورددراقتصادسنجی جدیدمی باشد.به هرحال ،قصدبراین است که این تکنیك راتوسعه دهیم وآنرادرموردبرآوردرویه ساده تمرکزیاتابع منحنی لورنزکه کاربردهای فراوانی دراقتصادوسایرعلوم مرتبطه داردبکاربریم .

درقمست ۲ موضوع راباتعریف نرم L_1 توابع پیوسته ومشخص کردن برآوردیا برازش یاتقریب تابع‌هدف رگرسیون باجمله‌تصادفی وپارامترهای مجهول مرتبط ادامه می دهیم . درقسمت های بعد،مسئله برآوردنرم L_1 الگوی خطی یك ودوپارامتری رادر حالت پیوسته حل خواهیم نمود . قبل ازاینکه این تکنیك هارادرموردبرآوردمنحنی لورنز بکاربریم مناسب است که نگاهی به نظریه رویه‌های تمرکزداشته باشیم .براین اساس ، در قسمت ۵ تعریف رویه تمرکزومخصوصا " منحنی لورنزارائه خواهدشدوفرم های تابعی پیشنهادشده برای آن موردبحث قرارخواهدگرفت . قسمت ۶ به تقریب نرم L_1 پیوسته منحنی لورنزاختصاص‌داده میشود .دراین قسمت ،باروشن ساختن اینکه چگونه می توان اطلاعات‌تابع‌توزیع احتمالی درآمدرابرای تعیین دقیق تقریب نرم L_1 منحنی لورنزمربوطه برای جامعه‌آماری تحت مطالعه‌بکارگرفت موضوع راادامه‌می دهیم .

۲ـ نرم L_1 توابع‌پیوسته

بطورکلی ،نرم L_p هرتابع $f(x)$ (نگاه‌کنیدبه (1964) Rice ,White) به شکل زیرتعریف، می شود .

$$||f(x)||_p = \left(\int_{x\in I} |f(x)|^p dx\right)^{1/p} \qquad (1)$$

که I یك مجموعه بسته کراندارمی باشد.نرم L_1 $f(x)$ به‌سادگی به‌شکل زیرنوشته می شود .

$$||f(x)||_1 = \int_{x\in I} |f(x)|dx \quad (۲)$$

فرض کنیدتابع غیرتصادفی x ، $f(x,\beta)$ باجمله‌تصادفی u ترکیب شده تا $y(x)$ رابه‌شکل زیرتشکیل دهد،

$$y(x) = f(x,\beta) + u \quad (۳)$$

که β بردارپارامترهای مجهول تابع f می باشد. بانوشتن u بعنوان پسماند $y(x)-f(x,\beta)$ ، در تقریب نرم L_1 بایدمقداربردار β راچنان تعیین کنیم که نرم L_1 جمله u حداقل شود. به عبارت دیگر،

$$\min_{\beta}: S=||u||_1=||y(x)-f(x,\beta)||_1=\int_{x\in I} |y(x)-f(x,\beta)|dx \quad (۴)$$

۳- برارزش نرم L_1 یك پارامتری خطی

مجددا" $f(x,\beta)$ رامانندتابع خطی زیرتعریف می کنیم،

$$y(x) = \beta x + u \quad (۵)$$

عبارت (۴) خلاصه می شودبه

$$\min_{\beta}: S = ||u||_1 = ||y(x)-\beta x||_1 = \int_{x\in I} |y(x)-\beta x|dx \quad (۶)$$

درفصل سوم مشابه گسسته (۶) توضیح داده شدوبا ارجاع به روش میانه‌وزنی لاپلاس ، محاسبه مقداربهینه β تشریح گردید. دریك روش دیگراستفاده مشتق های گسسته و معمولی بامتغیرکمکی t به عنوان نقطه تمییزپسماندهای منفی ومثبت قبل ازهرگونه محاسبه پیشنهادگردید. درحداقل کردن (۶) یك روشی مشابه‌رادوباره بکارخواهیم بست. بدین منظوردراین حالت شرایط لیپشیتزبایددرتوابع موردنظرصادق باشد(نگاه کنیدبه (1967-a) Usow و (1963) Dieudonne) عبارت (۶۶) رابه شکل زیرمی‌نویسیم

$$\min_{\beta}: S = \int_{x\in I} |x||y(x)/x - \beta|dx \quad (۷)$$

برای سهولت I رابه عنوان فاصله بسته $[0,1]$ تعریف می کنیم . روش مورد نظرمی‌تواندبه‌سایرفواصل دیگربدون هیچ اشکال عمده‌ای تعمیم داده شود(نگاه کنیدبه

Rivlin (1965) Hobby . Rice (1965)، Usow (1967 a) و Kripke) ازآنجائیکه x متعلق به فاصله بسته I می باشد، y(x) که یك تابع خطی از x و همچنین x/(x)y توابع هموار وپیوسته خواهندبود ،بدین ترتیب ازآنجائیکه y(x)/x تابع افزایشی یاکاهشی یکنواخت از x است ،یك مقدار t ∈ I رامی توان پیداکردکه خواص زیررادارا باشد.

$$\begin{aligned} y(x)/x &< \beta & \text{if } x &< t \\ y(x)/x &= \beta & \text{if } x &= t \\ y(x)/x &> \beta & \text{if } x &> t \end{aligned} \qquad (۸)$$

براساس (۸) می توان (۷) رابعنوان دوانتگرال معین مجزا با حدودبالا وپائین متفاوت نوشت :

$$\min_{\beta}: S = -\int_0^t |x|(y(x)/x - \beta)dx + \int_t^1 |x|(y(x)/x - \beta)dx \qquad (۹)$$

برای حل (۹) ، مشتقات جزئی آنرانسبت به t و β محاسبه می نمائیم

$$\frac{\delta S}{\delta \beta} = \int_0^t |x|dx - \int_t^1 |x|dx = 0 \qquad (۱۰)$$

وبااستفاده ازقضیه لایبنیتزبرای مشتق گیری انتگرال هانیست به حدودهای متغیرآنان t، داریم ،

$$\frac{\delta S}{\delta t} = -|t|[\frac{y(t)}{t} - \beta] - |t|[\frac{y(t)}{t} - \beta] = 0 \qquad (۱۱)$$

ازآنجائیکه x متعلق به [0,1] است ،معادله (۱۰) رامی توان به شکل زیر نوشت ،

$$\int_0^t x dx - \int_t^1 x dx = 0 \qquad (۱۲)$$

یا

$$\tfrac{1}{2} x^2 \Big]_0^t - \tfrac{1}{2} x^2 \Big]_t^1 = 0 \qquad (۱۳)$$

یا

$$\tfrac{1}{2} t^2 - \tfrac{1}{2} + \tfrac{1}{2} t^2 = 0 \qquad (۱۴)$$

نتیجه می دهد،

$$t = \sqrt{2}/2 \qquad (۱۵)$$

باجایگزینی در t درمعادله (۱۱) داریم

$$-2|\sqrt{2}/2|\,[\frac{y(\sqrt{2}/2)}{\sqrt{2}/2} - \beta] = 0 \qquad (۱۶)$$

یا

$$\beta = \frac{y(\sqrt{2}/2)}{\sqrt{2}/2} \qquad (۱۷)$$

بیادداشته باشیدکه $y(t)$ مقدارتابع $y(x)$ در $x=t$ می باشد. مقدار β درمعادله (۱۷) جواب بهینه (۶) می باشد. روش با لاعملا" بدست آوردن میانه وزنی درحالت پیوسته است .

قبل ازبحث درموردکاربرداین روش درمنحنی لورنز ،اجازه دهیدکه آنرابرای برآورد نرم L_1 الگوی خطی دوپارامتری برای داده های پیوسته بسط دهیم .

۴- برازش نرم L_1 دوپارامتری خطی

حال ،سعی براین است که تکنیك فوق رابرای الگوی خطی دوپارامتری بکاوبریم. (۴) رابه شکل زیرمی نویسم ،

$$\min_{\alpha,\beta}: S=||u||_1=||y(x)-\alpha-\beta x||_1=\int_{x\in I} |y(x)-\alpha-\beta x|\,dx \qquad (۱۸)$$

براساس (Rice 1964 c) ،اگر $f(\alpha^*,\beta^*,x)$ ، $y(x)$ را درمجموعه نقاط کانونی $\{x_i\,;i=1,2\}$ ، درون یابی نماید ،اگر $y(x)$ چنین باشدکه $y(x)-f(\alpha^*,\beta^*,x)$ فقط دراین x_i هاتغییرعلامت دهدونه درنقاط دیگر $[0,1]$ ، $f(\alpha^*,\beta^*,x)$ بهترین تقریب نرم L_1 $y(x)$ می باشد (همچنین نگاه کنیدبه (Usow 1967a) . بااستفاده ازاین قاعده ،اگراین نقاط را با t_1 و t_2 مشخص کنیم می‌توانیم (۱۸) رابرای $I=[0,1]$ به شکل زیر بنویسیم،

$$S = \int_0^{t_1} [y(x)-\alpha-\beta x]dx - \int_{t_1}^{t_2} [y(x)-\alpha-\beta x]dx + \int_{t_2}^{1} [y(x)-\alpha-\beta x]dx \qquad (۱۹)$$

ازآنجائیکه t_1 و t_2 نیزمجهول هستند،باید S راباتوجه به β ، α، t_1، t_2 حداقل کنیم ۰ مشتقات جزئی (۱۹) رابااستفاده ازقاعده لابینیتز نسبت به‌این متغیرها بدست آورده ومساوی صفرقرارمی دهیم ،داریم ،

$$\frac{\delta S}{\delta \alpha} = - \int_0^{t_1} dx + \int_{t_1}^{t_2} dx - \int_{t_2}^{1} dx = 0 \qquad (۲۰)$$

$$\frac{\delta S}{\delta \beta} = - \int_0^{t_1} xdx + \int_{t_1}^{t_2} xdx - \int_{t_2}^{1} xdx = 0 \qquad (۲۱)$$

$$\frac{\delta S}{\delta t_1} = 2[y(t_1) - \alpha - \beta t_1] = 0 \qquad (۲۲)$$

$$\frac{\delta S}{\delta t_2} = - 2[y(t_2) - \alpha - \beta t_2] = 0 \qquad (۲۳)$$

معادلات (۲۰) تا (۲۳) رامیتوان برای مقادیر α، β، t_1، t_2 بطور همزمان حل کردبنابراین ،سیستم معادلات زیرراداریم ۰

$$2t_2 - 2t_1 - 1 = 0 \qquad (۲۴)$$

$$t_2^2 - t_1^2 - ½ = 0 \qquad (۲۵)$$

$$y(t_1) - \alpha - \beta t_1 = 0 \qquad (۲۶)$$

$$y(t_2) - \alpha - \beta t_2 = 0 \qquad (۲۷)$$

جواب برابراست با

$$t_1=1/4 \qquad (۲۸)$$

$$t_2=3/4 \qquad (۲۹)$$

$$\alpha = y(3/4)-(3/4)\beta = y(1/4)-(1/4)\beta \qquad (۳۰)$$

$$\beta = 2[y(3/4)-y(1/4)] \qquad (۳۱)$$

قدم های فوق برای رگرسیون های ساده نرم L_1 پیوسته باقیدوبدون قید ،بعدا "برای محاسبه برآوردنرم L_1 منحنی لورنزبااستفاده ازتوزیع احتمالی درآمدبکارگرفته

خواهدشد.

این روش ،مشابه الگوی رگرسیون چندمتغیره برای حالت گسسته رامیتوان بسط دادکه m پارامترمجهول رادربرگیرد.به هرحال ،دراینجادراین بحث واردنمی شویم وآنرا به تحقیقات بعدی واگذارمی‌نمائیم .بعضی روشهای محاسباتی برای حل حالت مختلف الگوی m پارامتری توسط (1958) Ptak ، (1964) Rice ,white (1967a) Usow ، Lazarski(1975a,b,c, 77) بررسی می شوند. (همچنین نگاه کنیدبه ،(1965) Hobby, Rice ، (1965) Rivlin ، (1981) Kripke ، Watson) . حال نگاهی می اندازیم بررویه‌های تمرکزواشکال تابعی پیشنهادی برای حالت ساده آن یامنحنی لورنز.

۵- رویه تمرکز

تجزیه وتحلیل های تمرکزروشی است برای بررسی توزیع‌های چوله .ازآنجائیکه توزیع‌های چوله درون مقوله توزیع های غیرنرمال قرارمی گیرند ،ارتباط آن باآمارنیرومند ونرم L_1 افزایش می یابد.توزیع درآمدا غلب ازتوزیع‌های نظیرلوگ- نرمال یاپارتو (نگاه کنیدبه ،(1973) Cramer و (1989) Bidabad) تبعیت می کندوازاین بابت بارویه تمرکزمرتبط می شود.

رویه ساده تمرکزیا منحنی لورنزبرای یک متغیرتصادفی باتابع چگالی احتمال $f(v)$ رامی توان به شکل جفت مرتب زیرتعریف نمود،

$$\left(P(V|V\leq v),\frac{E(V|V\leq v)}{E(V)}\right) \qquad v\in R \tag{۳۲}$$

(1972a,b,c, 73, 81, 83, 87,85) Taguchi عنصردوم (۳۲) را در $P(V|V\leq v)$ ضرب می نمایدکه صحیح نمی باشد.تعریف وی برای (۳۳) معادل تعریف مامی باشد.برای تابع چگالی پیوسته $f(v)$ ، (۳۲) رامی توان به شکل زیرنوشت

$$\left(\int_{-\infty}^{v} f(w)dw,\frac{\int_{-\infty}^{v} wf(w)dw}{\int_{-\infty}^{+\infty} wf(w)dw}\right) \equiv (y(x(v)),x(v)) \tag{۳۳}$$

(۳۳) رابا $(y(x(v)),x(v))$ مشخص می کنیم که $x(v)$ و $y(x(v))$ عناصرآن هستند. بنابراین، y تابعی است که $x(v)$ رابه $y(x(v))$ می برد و $x(v)$ تابعی است که v رابه $x(v)$ می برد. تابع $y(x)$ تابع منحنی لورنزمی باشد. برای بحث درخواص منحنی لورنزرجوع کنیدبه (Kendall , Stuart 1977)

رویه تمرکزبه حالات دووچندمتغیره ،بسط داده شدکه دراینجاراجع به آن صحبت نخواهدشد. برای اطلاعات بیشتررجوع شودبه (Taguchi 1972a, b,c, 73,81,83,87, 88) . به هرحال دراین متن تقریب تابع منحنی لورنزبا استفاده ازنرم L_1 مورد نظر است که کاربردهای زیادی دراقتصاددارد.

توزیع درآمدا غلب بروی یك منحنی لورنز تصویرمی شود. درسالهای اخیرچندفرم تابعی برای آن معرفی شده است . این فرم هابایدخواصی راداشته باشندوهمچنین براحتی بتوان آنان رابا روش های برآورد خوب برآوردنمود. (Gupta 1984) خواص زیر را تعریف نمودیك تابع $f(x)$ منحنی لورنزرابیان می کنداگر :

$$
\begin{array}{lll}
\text{(الف)} & f(0)=0 & \\
\text{(ب)} & f(1)=1 & \\
\text{(ج)} & f'(x)\ge 0 \quad \text{for} \quad 0\le x<1 & \\
\text{(د)} & f''(x)\ge 0 \quad \text{for} \quad 0\le x\le 1 & \text{(۳۴)}\\
\text{(هـ)} & f(x)\le x \quad \text{for} \quad 0<x<1 & \\
\text{(و)} & 0\le\int_0^1 f(x)dx\le \tfrac{1}{2} &
\end{array}
$$

واضح است که هروقت (الف) تا (هـ) برقرارباشند، خاصیت (و) اضافه می باشد. زیرابا دست کاری کردن (الف) تا (هـ) داریم $0\le f(x)\le x$. باانتگرال گیری ازاین معادله (و) بدست می آید،

$$\int_0^1 0dx \le \int_0^1 f(x)dx \le \int_0^1 xdx \qquad (۳۵)$$

یا

$$0 \le \int_0^1 f(x)dx \le \tfrac{1}{2} \qquad (۳۶)$$

بنابراین خاصیت (و) همیشه برقراراست ، واحتیاج نداریم که آن رابرای توابع مورد

نظرکه شرایط (الف) تا (هـ) رابرقرارمی سازندآزمایش کنیم ۰

(Bidabad, Bidabad (1989a,b بردوخصیصه دیگرمنحنی لورنزکه از آن غفلت شده بودتاکیدمی نمایند ۰ اول ،منحنی لورنزمیتواندنسبت به خط $y=1-x$ برای $0\leq x\leq 1$ نامتقارن باشد ۰این موضوع سبب می شود ۰که منحنی های لورنزمختلف که دارای فرم تابعی یکسان باپارامترهای مختلف می باشندهمدیگررادر $0<x<1$ قطع کنند ۰

حال ، فرم های تابعی پیشنهادشده توسط سایرمحققین رابررسی می کنیم (1976) Podder و Kakwani تابع زیرراپیشنهادمی کند؛

$$M = aN^{l}(\sqrt{2}-N)^{c} \qquad (۳۷)$$

که $M=(x-y)/\sqrt{2}$ و $N=(x+y)/\sqrt{2}$ و $a\geq 0$ و $0\leq c\leq 1$ و $0\leq l\leq 1$ ۰ این فرم تمام خواص (۳۴) رابرقرارنمیسازد

(1980) Rasche, Gaffney ,Koo, Obst تابع زیرراپیشنهادمی کنند

$$y=[1-(1-x)^{a}]^{1/l} \qquad (۳۸)$$

که درآن $0\leq a\leq 1,\ 0\leq l\leq 1$ ۰ دراین فرم برآوردپارامترهاباروش حداقل مربعات مشکل می باشد ۰

(1984) Gupta فرم تابعی

$$y=xA^{x-1} \qquad (۲۹)$$

راپیشنهاد می نمایدکه $A>1$ ۰این فرم خواص مزبوررابرقرارمی سازدوبسهولت می تواندباروش حداقل مربعات برآوردشود ،ولی باتغییرپارامتر A (از A_i به A_j) توابع حاصل (y_i , y_j) هیچگاه درفاصله $0<x<1$ متقاطع نخواهندبود ۰ برای اثبات این ،می توانیم دستگاه زیررابرای x و y حل کنیم ۰

$$\begin{aligned} y&=xA_i^{x-1} \\ y&=xA_j^{x-1} \end{aligned} \qquad (۴۰)$$

جوابها $x=y=0$ و $x=y=1$ هستندکه دردامنه $0<x<1$ قرارنمیگیرند۰

(1989a,b) Bidabad و Bidabad فرم تابعی زیررادرنظر می گیرندکه خواص (الف) تا (هـ) رابرقرارنموده وباتغییرپارامترهای آن ،منحنی های حاصل میتوانندهمدیگرراقطع نمایند.

$$y=x^{B}A^{x-1} \tag{۴۱}$$

که $B\geq 1,\ A\geq 1$ برای $0\leq x\leq 1$. خواص (۳۴) به شکل زیربرقرارمیشوند

$$\begin{aligned}
&(\text{الف}) && f(0)=0 \\
&(\text{ب}) && f(1)=1 \\
&(\text{ج}) && f'(x)=x^{B-1}A^{x-1}[B+x.\ln(A)]\geq 0 \quad \text{for } 0\leq x<1 \\
&(\text{د}) && f''(x)=x^{B-2}A^{x-1}\{[B+x.\ln(A)]^{2}-B\}\geq 0 \quad \text{for } 0\leq x\leq 1 \\
&(\text{هـ}) && f(x)=x^{B}A^{x-1}=(x^{B}/A^{1-x})\leq x \quad \text{for } 0<x<1
\end{aligned} \tag{۴۲}$$

اشکال مختلف تابع مانند $y=x^{B_i}A_i^{x-1}$ و $y=x^{B_j}A_j^{x-1}$ میتوانند درفاصله $0<x<1$ متقاطع باشند .با حل همزمان معادلات زیر ،مکان هندسی نقـــاط تقاطع با مقادیر مختلف پارامترهای دومنحنی لورنزبدست می آید .

$$\begin{aligned} y&=x^{B_i}A_i^{x-1} \\ y&=x^{B_j}A_j^{x-1} \end{aligned} \tag{۴۳}$$

جواب برابراست با ،

$$\frac{x-1}{\log(x)}=\frac{B_j-B_i}{\log(A_i/A_j)} \tag{۴۴}$$

واضح است که زمانی که (۴۴) برقرارشود ،یك تقاطع دردومنحنی (۴۳) پیدا می شود .بنا- براین اگردستگاه زیرراحل کنیم ،

$$\begin{aligned} x-1&=B_j-B_i \\ \log x&=\log(A_i/A_j) \end{aligned} \tag{۴۵}$$

میتوان یك رابطه برحسب A_i , A_j , B_i , B_j که (۴۴) رابرقرارنماید بدست آوریم . بنابراین

$$A_i / A_j - 1 = B_j - B_i \qquad (۴۶)$$

بدین ترتیب هروقت A_i, A_j, B_i, B_j بتوانندمعادله (۴۶) رابرقرارنمایند، یك جواب (یا تقاطع) برای (۴۳) وجودخواهدداشت. ولی تقاطع دردرون دامنه $0 < x < 1$ است. اگر،

$$0 < x = B_j - B_i + 1 < 1$$
$$0 < x = (A_i / A_j) < 1 \qquad (۴۷)$$

یا،

$$0 < B_j - B_i < 1$$
$$0 < A_i < A_j \qquad (۴۸)$$

موضوع دومی که (Bidabad , Bidabad 1989 a,b) تاکیـــد می نمایندطبیعت خودهمبستگی خطاهادرداده‌های منحنی لورنزمی باشد، زمانیکه یك خطا در($t - 1$) امیـــــن درصدصاحبان درآمدباشد، این خطا عینا " به درصدانباشتــــه بعدی (t) منتقل می شود. این بدلیل این است که داده‌های انباشته برای برآوردمنحنی لورنزبکارگرفته می شود. بنابراین، اگر u_t رابعنوان جمله اخلال مشاهده t ام (درصدانباشته) تعریف کنیم، تشخیص خودهمبستگی خطاهاازقرارزیرخواهدبود،

$$u_t = u_{t-1} + w_t \qquad (۴۹)$$

که w_t ازفروض کلاسیك رگرسیون پیروی می کند. بنابراین فرم تصادفی فرم تابـــع پیشنهادی مارامی توان به شکل زیرنوشت،

$$y_t = x_t^{B} . A^{x_t - 1} . e^{u_t} \qquad (۵۰)$$

تابع (۵۰) رابرای یك زمان تاخیرمی نویسیم،

$$y_{t-1} = x_{t-1}^{B} . A^{x_{t-1} - 1} . e^{u_{t-1}} \qquad (۵۱)$$

طرفین (۵۰) رابرطرفین (۵۱) تقسیم کرده ولگاریتم طبیعی آنراحساب می کنیم،

$$\ln\left(\frac{y_t}{y_{t-1}}\right) = B . \ln\left(\frac{x_t}{x_{t-1}}\right) + \ln(A) . (x_t - x_{t-1}) + u_t - u_{t-1} \qquad (۵۲)$$

ازآنجائیکه $u_t - u_{t-1}$ برابر w_t از(۴۹) می باشدمسئله خودهمبستگی

وجودنداردو(۵۲) می توانندبا روش حداقل مربعات برآوردشود.

به هرحال ،برای برآوردکردن توابع فوق بایدداده‌هائی منفردراازجامعه‌آماری موردنظر جمع‌آوری نموده وپس ازمحاسبات لازم وایجادیک مجموعه اطلاعاتی مناسب خوراک روشهای برآوردراتهیه نمائیم . همانطورکه دربالا ذکرشد،خطاهای زیادی درحین ایجاداطلاعات بصورت انباشتی وسایرعملیات محاسباتی لازم واردخواهندشد.برای اجتناب ازاین مسائل بایدروش دیگری رابرای برآوردمنحنی لورنزجامعه‌آماری خوداتخاذنمائیم به هرحال ،اگر توزیع احتمالی درآمدشناخته شده باشد،به جای جمع‌آوری مشاهدات منفرد،می توانیم منحنی لورنزرابااستفاده‌ازروش برازش نرم L_1 توابع پیوسته برآوردنمائیم.درقسمت‌های بعداین روش رابرای برآوردپارامترهای الگوهای (۳۹) و (۴۱) بااستفاده ازاطلاعات تابع چگالی احتمال درآمدبکارخواهیم گرفت .

۶ـ برازش نرم L_1 پیوسته منحنی لورنز

برای برآوردکردن پارامترهای منحنی لورنززمانی که اطلاعات تابع چگالی احتمال درآمدداده شده‌است همیشه یک راههموار درپیش نداریم . زمانیکه تابع چگالی احتمال انتگرال‌پذیراست مشکل اساسی درپیش نیست. می توانیم رابطه تابعی بین دوعنصر (۳۳) رابامحاسبات ساده ریاضی بدست آوریم . ولی وقتی که انتگرال‌های (۳۳) قابل حصول نیستند،روش دیگری رابایدجستجونمود.

فرض کنیددرآمدیک جامعه باتابع چگالی احتمال $f(w)$ توزیع‌شده باشد.این تابع چگالی می‌تواندیک تابع چوله مانندپارتویالوگ ـ نرمال به صورت زیرباشد.

$$f(w)=\theta k^{\theta} w^{-\theta-1} \qquad w\geq k>0,\ \theta>0 \tag{۵۳}$$

$$f(w)=[1/w\sigma\sqrt{(2\pi)}]\exp\{-[\ln(w)-\mu]^2/2\sigma^2\}$$
$$w\in(0,\infty),\ \mu\in(-\infty,+\infty),\ \sigma>0 \tag{۵۴}$$

این دوتوزیع بعنوان نامزدهای خوبی برای توضیح درآمدفردی شناخته شده اند(نگاه کنیدبه (1973) Cramer , (Bidabad (1989)

درحالت تابع چگالی پارتو(۵۳) تابع‌منحنی لورنزرابه راحتی می توان به شکل زیر

بدست آورد. (w) F تابع توزیع پارتو

$$F(w)=1-(k/w)^{\theta} \tag{۵۵}$$

با میانگین زیر را نشان می دهد.

$$E(w)=\theta k/(\theta-1), \quad \theta>1 \tag{۵۶}$$

اگر تابع y ، که توسط (۳۳) بیان شده است را به عنوان تابعی از x پیدا کنیم تابع لورنز بدست خواهد آمد. حال ،به شکل زیر عمل می کنیم. جملات (۳۳) را جابجا می کنیم

$$y(x(v))=\int_{-\infty}^{v} f(w)dw \tag{۵۷}$$

$$x(v)=[1/E(x)]\int_{-\infty}^{v} wf(w)dw \tag{۵۸}$$

تابع توزیع پارتو را جایگزین می کنیم

$$y(x(v))=F(v)=1-(k/v)^{\theta} \tag{۵۹}$$

$$x(v)=[(\theta-1)/\theta k]\int_{k}^{v} w\theta k^{\theta} w^{-\theta-1} dw \tag{۶۰}$$

یا

$$x(v)=1-(k/v)^{\theta-1} \tag{۶۱}$$

حال با حل (۶۱) برای v وجایگزینی آن در (۵۹) ،منحنی لورنز برای توزیع پارتو به شکل زیر بدست می آید.

$$y=1-(1-x)^{\theta/(\theta-1)} \tag{۶۲}$$

واضح است که (۶۲) خواص منحنی لورنز که در (۳۴) بیان شده است را برقرار می سازد همانطور که درحالت توزیع پارتو (۵۳) نشان داده شد پیدا کردن منحنی لورنز به سادگی حاصل می شود. ولی ،اگر تابع چگالی لوگ - نرمال (۵۴) را انتخاب کنیم. روش ما مانند فوق نمی تواند باشد. زیرا انتگرال تابع لوگ - نرمال هنوز پیدا نشده است. بدین ترتیب روش دیگری برای حل این مشکل باید ابداع کنیم. یك راه استفاده از تقریب یا برازش نرم L_p میباشد. درصفحات پائین ،تکنیك برازش نرم L_1 برای برآورد پارامترهای فرم های معلوم تابعی (۳۹) و (۴۱) با استفاده از تابع چگالی احتمال پیوسته بکار گرفته خواهد شد. حال به شکل زیر عمل می کنیم.

براساس (۳۲) و(۳۳) متغیرهای تابع ومستقل (۳۹) و (۴۱) به شکل زیرنوشته می شوند،

$$y(x(v))=\int_0^v f(w)dw \tag{۶۳}$$

$$x(v)=[1/E(x)]\int_0^v wf(w)dw \tag{۶۴}$$

(۶۲) و(۶۴) رادرون (۳۹) جایگزین نموده وباتعریف جمله خطای تصادفی u داریم ،

$$\int_0^v f(w)dw=[1/E(w)]\int_0^v wf(w)dw.A^{[1/E(w)]\int_0^v wf(w)dw-1}e^{u} \tag{۶۵}$$

یابطورخلاصه

$$y(x)=xA^{x-1}e^{u} \tag{۶۶}$$

همینطوربرای الگوی (۴۱)،

$$\int_0^v f(w)dw=\{[1/E(w)]\int_0^v wf(w)dw\}^{B}A^{[1/E(w)]\int_0^v wf(w)dw-1}e^{u} \tag{۶۷}$$

یابه طورخلاصه

$$y(x)=x^{B}A^{x-1}e^{u} \tag{۶۸}$$

از (۶۶) و (۶۸) لگاریتم طبیعی می گیریم

$$\ln y(x)=\ln x + (x-1)\ln A + u \tag{۶۹}$$

$$\ln y(x)=B.\ln x + (x-1)\ln A + u \tag{۷۰}$$

باتوجه به خواص الف تا هـ در (۳۴) وتابع چگالی احتمال $f(w)$ ومعادلات (۶۵) تا(۶۸) ،واضح است که x متعلق به فاصله $[0,1]$ می باشد. بنابراین تابع ابژکتیو نرم L_1 برای حداقل کردن (۶۹) یا (۷۰) ازقرارزیراست .

$$\min:\ S = \int_0^1 |u|dx \tag{۷۱}$$

حال ،اجازه دهیدکه برآوردنرم L_1 پارامتر A برای فرم تابعی منحنی لورنـــــز که توسط (1984) Gupta پیشنهادشده ودر (۶۹) مجددا" تعریف گردیـــــده را بدست آوریم . تابع ابژکتیونرم L_1 آن ازقرارذیل است ،

$$\min_{A}:\ S = \int_0^1 |\ln y(x) - \ln x - (x-1)\ln A|dx \tag{۷۲}$$

یا

$$\min_{A}: S = \int_0^1 |x-1|\,|[\ln y(x)-\ln x]/(x-1) - \ln A|\,dx \qquad (۷۳)$$

بایک روش مشابه که درمورد (۹) استفاده شد، میتوانیم (۷۳) رابه شکل زیربنویسیم

$$\min_{A}: S = \int_0^t |x-1|\{[\ln y(x)-\ln x]/(x-1) - \ln A\}dx$$
$$- \int_t^1 |x-1|\{[\ln y(x)-\ln x]/(x-1) - \ln A\}dx \qquad (۷۴)$$

ازآنجائیکه $0\le x\le 1$ داریم ،

$$\min_{A}: S = \int_0^t [\ln y(x) - \ln x - (x-1)\ln A]dx$$
$$- \int_t^1 [\ln y(x) - \ln x - (x-1)\ln A]dx \qquad (۷۵)$$

مشتقات جزئی (۷۵) رانسبت به t و A حساب کرده ومساوی صفرقرارمی دهیم ،

$$\frac{\delta S}{\delta A} = -\int_0^t [(x-1)/A]dx + \int_t^1 [(x-1)/A]dx = 0 \qquad (۷۶)$$

$$\frac{\delta S}{\delta t} = 2[\ln y(t) - \ln t - (t-1)\ln A] = 0 \qquad (۷۷)$$

ازمعادله (۷۶) داریم ،

$$-\left[\tfrac{1}{2}x^2-x\right]_0^t + \left[\tfrac{1}{2}x^2-x\right]_t^1 = 0 \qquad (۷۸)$$

یا

$$t = 1\pm\sqrt{2}/2 \qquad (۷۹)$$

ازآنجائیکه t بایدمتعلق به فاصله $[0,1]$ باشدقبول می کنیم ،

$$t = 1-\sqrt{2}/2 \qquad (۸۰)$$

(۸۰) رادر (۷۷) جایگزین نموده

$$\ln y(1-\sqrt{2}/2) - \ln(1-\sqrt{2}/2) - (1-\sqrt{2}/2-1)\ln A = 0 \qquad (۸۱)$$

پس برآوردنرم L_1 A برابرخواهدبودبا

$$A = \left[\frac{1-\sqrt{2}/2}{y(1-\sqrt{2}/2)}\right]^{\sqrt{2}} \qquad (۸۲)$$

حال ،روش تقریب نرم L_1 رابرای فرم تابعی منحنی لورنز (۴۱) پیشنهادشده توسط (Bidabad , Bidabad 1989a,b) باتعریف مجدد(۷۰) بکار می بندیم . تابع ابژکتیونرم L_1 (۷۱) برای الگوی (۷۰) رامی نویسم ،

$$\min_{A,B}: S = \int_0^1 |\ln y(x) - B \ln x - (x-1) \ln A| dx \quad (۸۳)$$

یا

$$\min_{A,B}: S=\int_0^1 |x-1| |[\ln y(x)]/(x-1)-(\ln x)/(x-1)-\ln A| dx \quad (۸۴)$$

تابع ابژکتیو(۸۴) ـ باتغییری بروی متغیرها ـ شبیه (۱۸) می باشد . بنابراین ، بـا روشی مشابه (۱۶) تا (۳۱) می توانیم S رابه شکل زیربنویسیم ،

$$\begin{aligned}\min_{A,B}: S = &\int_0^{t1} |x-1| \{[\ln y(x)]/(x-1)-(\ln x)/(x-1)-\ln A\} dx\\ &- \int_{t1}^{t2} |x-1| \{[\ln y(x)]/(x-1)-(\ln x)/(x-1)-\ln A\} dx \\ &+ \int_{t1}^{1} |x-1| \{[\ln y(x)]/(x-1)-(\ln x)/(x-1)-\ln A\} dx\end{aligned} \quad (۸۵)$$

ازآنجائیکه $0 \le x \le 1$ (۸۵) خلاصه می شودبه ،

$$\begin{aligned}\min_{A,B}: S = &- \int_0^{t1} [\ln y(x) - B \ln x - (x-1) \ln A] dx\\ &+ \int_{t1}^{t2} [\ln y(x) - B \ln x - (x-1) \ln A] dx\\ &- \int_{t2}^{1} [\ln y(x) - B \ln x - (x-1) \ln A] dx\end{aligned} \quad (۸۶)$$

مشتقات جزئی S رانسبت به A ، B ، t_1 ، t_2 پیداکرده ومساوی صفر قرارمی دهیم .

$$\frac{\delta S}{\delta A} = \frac{1}{A}\left[\int_0^{t_1}(x-1)dx - \int_{t_1}^{t_2}(x-1)dx + \int_{t_2}^{1}(x-1)dx\right] = 0 \quad (۸۷)$$

$$\frac{\delta S}{\delta B} = \int_0^{t_1}\ln(x)dx - \int_{t_1}^{t_2}\ln(x)dx + \int_{t_2}^{1}\ln(x)dx\Big] = 0 \quad (۸۸)$$

$$\frac{\delta S}{\delta t_1} = -2\{\ln[y(t_1)] - B\ln(t_1) - (t_1-1)\ln(A)\} = 0 \quad (۸۹)$$

$$\frac{\delta S}{\delta t_2} = 2\{\ln[y(t_2)] - B\ln(t_2) - (t_2-1)\ln(A)\} = 0 \quad (۹۰)$$

سیستم معادلات همزمان فوق را میتوان برای مجهولات t_1 , t_2, A , B حل نمود.

معادله (۸۷) خلاصه می شود به ،

$$t_1^2-t_2^2-2(t_1-t_2)-\tfrac{1}{2} = 0 \qquad (۹۱)$$

معادله (۸۸) را میتوان به شکل زیر نوشت

$$[\, x(\ln x-1)\,]_0^{t_1} - [\, x(\ln x-1)\,]_{t_1}^{t_2} - [\, x(\ln x-1)\,]_{t_2}^{1} = 0 \qquad (۹۲)$$

یا

$$t_1(\ln t_1-1)-t_2(\ln t_2-1)+t_1(\ln t_1-1)+\ln 1-1-t_2(\ln t_2-1)=0 \qquad (۹۳)$$

یا

$$t_1(\ln t_1-1) - t_2(\ln t_2-1) - \tfrac{1}{2} = 0 \qquad (۹۴)$$

t_1 را از (۹۱) حساب می کنیم ،

$$t_1 = 1 \pm \sqrt{(t_2^2-2t_2+3/2)} \qquad (۹۵)$$

چون $0 \le t_1 \le 1$ است رابطه زیر را قبول می کنیم

$$t_1 = 1 - \sqrt{(t_2^2-2t_2+3/2)} \qquad (۹۶)$$

t_1 از (۹۶) را در (۹۴) جایگزین می کنیم

$$[1 - \sqrt{(t_2^2-2t_2+3/2)}]\{\ln[1 - \sqrt{(t_2^2-2t_2+3/2)}] - 1\}$$
$$- t_2(\ln t_2-1) - \tfrac{1}{2} = 0 \qquad (۹۷)$$

جملات را جابجا می کنیم

$$\ln \frac{[1-\sqrt{(t_2^2-2t_2+3/2)}]^{[1-\sqrt{(t_2^2-2t_2+3/2)}]}}{t_2^{t_2}} + t_2 - 3/2$$
$$+ \sqrt{(t_2^2-2t_2+3/2)} = 0 \qquad (۹۸)$$

ریشه معادله (۹۸) را میتوان با یک الگوریتم عددی مناسب پیدا نمود. مقدار آن تا پنج

رقم گردشده برابراست با

$$t_2 = 0.40442 \tag{۹۹}$$

مقدار t_1 باجایگزینی t_2 در(۹۶) بدست می آید۰

$$t_1 = 0.07549 \tag{۱۰۰}$$

بااستفاده از t_2 , t_1 در(۹۹) و (۱۰۰) مقادیر A, B رااز(۸۹) و(۹۰) بدست می آوریم

$$B = \frac{(t_2-1)\ln y(t_1) - (t_1-1)\ln y(t_2)}{(t_2-1)\ln(t_1) - (t_1-1)\ln(t_2)} \tag{۱۰۱}$$

یا

$$B = -0.84857\ln[y(0.07549)] + 1.31722\ln[y(0.40442)] \tag{۱۰۲}$$

و

$$A = [y(0.07549)]^{-1.28986}[y(0.40442)]^{3.68126} \tag{۱۰۳}$$

حال اجازه دهیدکه ببینیم چگونه معادله (۸۲) برای الگوی (۳۹) ومعادلات (۱۰۲)و(۱۰۳) برای الگوی (۴۱) میتوانندمورداستفاده واقع شوندتاپارامترهای منحنی لورنزرابااستفاده ازتابع‌توزیع احتمال بدست آورد،درالگوی (۳۹) باید (۶۴) رابرای $x(v)=1-\sqrt{2}/2$ حل کنیم۰ به عبارت دیگر،مقدار v رابایدچنان پیداکردکه

$$x(v) = [1/E(w)]\int_0^v wf(w)dw = 1-\sqrt{2}/2 \tag{۱۰۴}$$

باجایگزینی این مقدار v درون(۶۳)، مقدار $y(1-\sqrt{2}/2)$ محاسبه می شود۰ مقدار $y(1-\sqrt{2}/2)$ برای محاسبه پارامتر A در(۸۲) برای الگوی (۳۹) بکار برده می شود۰ روش مزبوربرای الگوی (۴۱) مشابه (۳۹) می باشدبااین اختلاف،که دومقدار v رابایدحساب کرد۰ دومقدارمختلف، v به شکل زیرمحاسبه می شوند

$$x(v) = [1/E(w)]\int_0^v wf(w)dw = 0.07549 \tag{۱۰۵}$$

$$x(v) = [1/E(w)]\int_0^v wf(w)dw = 0.40442 \tag{۱۰۶}$$

مقادیر v رادر (۶۲) جایگزین نموده تا مقادیر $y(0.07549)$ و $y(0.40442)$ را محاسبه نمائیم . این مقادیر y برای محاسبه پارامترهای الگوی (۴۱) با جایگزین کردن آنها در (۱۰۲) و (۱۰۳) بکار گرفته می شوند .

تنها مسئله ای که باقی می ماند محاسبه انتگرالهای معین $x(v)$ که در (۱۰۴) ، (۱۰۵)،(۱۰۶) و (۶۳) تعریف شده اند می باشد که میتوان با استفاده از روشهای عددی مناسب آنها را بدست آورد .

فصل ششم

"خلاصه ونتایج ،توصیه برای تحقیقات بعدی"

۱۰ خلاصه ونتایج

درفصل اول سعی براین داشتیم که خواننده رابامفهوم واهمیت برآوردکننده نرم L_1 آشنانمائیم . بحث شدکه چرابرآوردکننده حداقل مربعات درحالات توزیع خطای Gaussian non- قابل اتکانیست .بانگاهی به تاریخ شماری تحلیل های آماری برمبنای نرم L_1 به مسائل مربوط به این برآوردکننده اشاره شدوهدف این تحقیق بیان گردید . قدم های اصلی که درفصول بعدی میبایست برداشته می شدنیزذکرگردید. تعریف وخواص فضای نرم بطور خلاصه مرورشدتاخواننده رابه علائم ریاضی سایرفصول آشناسازد . موقعیت برآوردهای نرم L_1 درخانواده برآوردهای نرم L_p ملاحظه گردیدوخواص برآوردنرم L_1 ارائه شد . ازاین بابت خاصیت تغییرناپذیری ،تبدیل برروی متغیرها ،تحدب تابع ابژکتیو ، پسماندهای صفردرجواب بهینه ،شرط بهینه بودن ،جواب های یکتاوغیریکتاوتحلیل های داخلی وحساسیت برآوردنرم L_1 همگی موردبحث قرارگرفتند .

درفصل دوم سعی براین داشتیم که یك مرورعمومی برادبیات معیارنرم L_1 درزمینه نظریه برآوردوموضوعات مربوطه باتوجه خاص به روش های محاسباتی داشته باشیم . از آنجائیکه درفصل سوم چندالگوریتم نزولی رابرای محاسبه برآوردنرم L_1 پارامترهای رگرسیون توسعه می دهیم ،بخشی که به این نوع الگوریتم هاتخصیص یافته تشریحی تر میباشد . به هرحال ،تاریخ شماری وتوسعه تاریخی نظریه برآوردنرم L_1 برای دوره ۱۹۲۸- ۱۶۳۲ بررسی شدوالگوریتم های متعلق به دوره بعداز ۱۹۲۸ به سه گروه عمده نزولی مستقیم ،نوع سیمپلکس وسایرالگوریتم هاطبقه بندی شدند . بحث هائی درموردمسئلهمقدار اولیه برای شروع الگوریتم های مختلف ،نرم افزارهای کامپیوتری موجود ،مقایسه الگوریتم هاکه قبلا" توسط سایرمحققین انجام شده ونگاهی کوتاه برروشهای محاسبه نرم غیرخطی وادبیات محاسبه نرم L_p همگی انجام گرفت . تشخیص نرم L_1 دستگاه معادلات همزمان که زمینه موردعلاقه اقتصادسنجان می باشدموربحث قرارگرفت . جنبه های آماری نرم L_1 همانند ،توزیع نمونه گیری ،استنباط آماری ،آمارچندمتغیره ،غیرپارامتری ونیرومندبطورخلاصه ذکرگردیدند . کاربردنرم L_1 درزمینههای مختلف علم باتاکیدبیشتر برمسائل اقتصادی موضوع دیگری درفصل دوم بود . درپایان این فصل این مروررابا ارائه توضیحی درموردسایرمشتقات نرم L_1 خاتمه دادیم .

درفصل سوم چندالگوریتم برای محاسبه نرم L_1 پارامترهای رگرسیون پیشنهادشدکه دوتا ازآنهاکه ازکارآئی بیشتری برخورداربودندبرای الگوهای رگرسیون ساده وچندمتغیره انتخاب شدند. به هرحال ،بارگرسیون خطی ساده باقیدوحل ومحاسبه مسئله میانه وزنی مربوطه مسئله راآغازنمودیم . دراین جهت یك تابع محاسباتی معرفی گردید. بابحث بر الگوی m پارامتری بحث راباتوسعه الگوریتم برای شمول رگرسیون خطی ساده بدون قید ادامه دادیم ودوالگوریتم خام وکارآ پیشنهادگردیدند. یك برنامه کامپیوتری نیزمعرفی شد. این روش سرانجام بامعرفی دوالگوریتم جدیدبه الگوی m پارامتری تعمیم یافت که الگوریتم ۴ بعنوان کارآترانتخاب گردید. خواص مختلف این الگوریتم هاموردبحث قرار گرفت وبرنامه‌های کامپیوتری مربوطه معرفی شدند. درقسمت آخرنگاهی گذرابرمسئله‌مقدار اولیه داشتیم .

ازآنجائیکه محاسبه جواب صحیح درالگوریتم های محاسباتی مهم ترازکارآئی می باشد آن الگوریتم هائی بایدانتخاب شوندکه جواب صحیح رامحاسبه می نمایندوازمیان آنها ، سریع‌ترین آنهاراانتخاب کرد.ازاین بابت ،الگوریتم ۲ والگوریتم Sposito (1983) Josvanger درفصل چهارم بایکدیگرمقایسه شدند. درتمام آزمایشات هر دو الگوریتم برای الگوی رگرسیون نرم L_1 خطی دوپارامتری جواب راصحیح محاسبه کردند. الگوریتم ۲ که سریعترازدیگری بود برای استفاده عملی معرفی گردید.برای رگرسیون چندمتغیره ،الگوریتم های Bloomfield, Steiger (1980) و Kung (1979) و Armstrong,Frome (1973) Barrodale, Roberts درمقایسه باالگوریتم ۴ قرارگرفتند. دوالگوریتم اول موفق به‌محاسبه جواب صحیح درالگوهای بزرگترنگردیدند. بدلیل انباشت خطای گردکردن این دوالگوریتم در مسئله های با اندازه بزرگ قابل استفاده نبودند. به‌هرحال نتیجه گیری می شودکه الگوریتم ۴ برای الگوهای باکمترازده‌پارامتر والگوریتم Barrodale , Roberts (1973,74) برای الگوی ده پارامتری مناسب ترمی باشند.این نتیجه آخرخیلی قطعی نیست ،زیرادرحالت الگوی ده‌پارامتری با ۱۰۰۰۰ مشاهده الگوریتم ۴ خیلی برترازالگوریتم Barrodale , Roderts (1973) بود. به‌هرحال ،ازآنجائیکه درکارهای کاربردی همیشه باالگوهای خیلی بزرگ مواجه نیستیم ،این نتیجه گیری ازلحاظ عملی ضعیف است .

فصل پنجم به توسعه تکنیک برآوردنرم L_1 برای حالت پیوسته تخصیص داده شد دوروش برای برازش نرم L_1 الگوهای خطی ساده بدون قیدوباقیدباتوابع خطای پیوسته پیشنهاد گردید. بانگاهی برنظریه رویه تمرکزوفرم های تابعی قبلا" معرفی شده برای منحنی لورنز برآوردپارامترهای آنان موردملاحظه قرارگرفت . با اشاره به دوتابع توزیع احتمال پارتو و لوگ - نرمال ، که مناسب الگوسازی درتوزیع درآمدمی باشد ،تکنیک های متفاوتی برای پیوستن آنها به منحنی لورنزتوسعه داده شدکه قابل استفاده برای سایرتوزیع های احتمال می باشد.

۲. توصیه برای تحقیقات بعدی

درسالهای اخیرمعیارنرم L_1 جذابیت زیادی درزمینه های متفاوت علوم پیداکرده است وبنابراین جلب نظرمحققین به این موضوع لازم بنظرمی رسد. بهرحال ،درفرآیند این مطالعه به نکاتی دست یافتیم که میتواندبه توسعه کاربردومحاسبه نرم L_1 بیانجامد.

۱) تکنیک های تشخیص دهنده مانندتحلیل های داخلی وحساسیت رگرسیون نرم L_1 احتیاج به بررسی وتوسعه بیشتردارند. برای بحث اساسی دراین موردبه (Narula ,Wellington 1985) وهمچنین فصل اول مراجعه نمائید.

۲) درالگوریتم پیشنهادی ۴ ،موفق به یافتن معیاری برای حذف آن مشاهده ازمجموعه مشاهدات موجوددرپایه که مقدارتابع ابژکتیورا ازسایرنقاط بیشترکاهش می دهد نشدیم. اگربتوان معیاری همانندروش اکتشافی (Steiger 1980) و Bloomfield (همچنین نگاه کنیدبه فصل دوم) پیداکرد ،تعداد دورهای تکرارخیلی کمترخواهدشدوزمان محاسبه راکوتاه ترمی نماید.

۳) مانع اصلی برسرراه بکارگیری تکنیک مشتق گیری گسسته به مسائل عملی کەدرفصل سوم پیشنهادشد ،مرتب کردن مجددپسماندها بطورنزولی وقبل ازشروع محاسبات می باشد. اگراین مانع رفع شود ،یك الگوریتم مشابه الگوریتم خطی باقیدساده را میتوان برای رگرسیون چندمتغیره پیشنهادنمود.

۴) الگوریتم های بهبودیافته جدیدبرای برنامەریزی خطی ،همانندروش تصویری Karmarkar امکان بکارگیری برای مسئله نرم L_1 رادارند.

۵) روش مشابه الگوی رگرسیون چندمتغیره برای حالت گسسته را میتوان برای شمول m پارامترمجهول روش برازش نرم L_1 پیوسته بسط داد. حالت های الگوهای یك ودوپارامتری درفصل پنجم توسعه یافتند.

" ضمائم "

برای توضیح الگوریتم به قسمت ۱۰۲ فصل سوم رجوع کنید.

```
      FUNCTION LWMED(N,YS,W,L)
      REAL YS(N),W(N)
      INTEGER L(N),HI
      II=0
      SHI=0.
      SLO=0.
      SZ=0.
      SP=0.
      SN=0.
      DO 4  I=1,N
      W(I)=ABS(W(I))
      IF(YS(I))3 ,2 ,1
 1    SP=SP+W(I)
      GO TO 4
 2    SZ=SZ+W(I)
      GO TO 4
 3    SN=SN+W(I)
 4    CONTINUE
      SHI=SP+SZ
      SLO=SN+SZ
      IF(SHI.LE.SLO)GO TO 6
      SHI=0.
      DO 5 I=1,N
      IF(YS(I).LE.0.) GO TO 5
      II=II+1
      L(II)=I
 5    CONTINUE
      GO TO 8
 6    SLO=0.0
      DO 7 I=1,N
      IF(YS(I).GT.0.)GO TO 7
      II=II+1
      L(II)=I
 7    CONTINUE
 8    LO=1
      HI=II
 10   IF(HI.GT.LO+1)GO TO 30
      LWMED=L(LO)
      IF(LO.EQ.HI) RETURN
      IF(YS(L(LO)).LE.YS(L(HI))) GO TO 20
      LT=L(LO)
      L(LO)=L(HI)
      L(HI)=LT
      LWMED=L(LO)
 20   IF(SHI+W(L(HI)).GT.SLO+W(L(LO))) LWMED=L(HI)
      RETURN
 30   MID=(LO+HI)/2
      LOP=LO+1
      LT=L(MID)
      L(MID)=L(LOP)
      L(LOP)=LT
      IF(YS(L(LOP)).LE.YS(L(HI))) GO TO 40
      LT=L(LOP)
      L(LOP)=L(HI)
      L(HI)=LT
 40   IF(YS(L(LO)).LE.YS(L(HI)))GO TO 50
      LT=L(LO)
      L(LO)=L(HI)
      L(HI)=LT
 50   IF(YS(L(LOP)).LE.YS(L(LO))) GO TO 60
      LT=L(LOP)
      L(LOP)=L(LO)
      L(LO)=LT
 60   LWMED=L(LO)
```

```
      I=LOP
      J=HI
      XT=YS(LWMED)
      TLO=SLO
      THI=SHI
70    TLO=TLO+W(L(I))
      I=I+1
      IF(YS(L(I)).LT.XT) GO TO 70
80    THI=THI+W(L(J))
      J=J-1
      IF(YS(L(J)).GT.XT) GO TO 80
      IF(J.LE.I) GO TO 90
      LT=L(I)
      L(I)=L(J)
      L(J)=LT
      GO TO 70
90    TEST=W(LWMED)
      IF(I.NE.J) GO TO 100
      TEST=TEST+W(L(I))
      I=I+1
      J=J-1
100   IF(TEST.GE.ABS(THI-TLO)) RETURN
      IF(TLO.GT.THI)GO TO 110
      SLO=TLO+TEST
      LO=I
      GO TO 10
110   SHI=THI+TEST
      LO=LOP
      HI=J
      GO TO 10
      END
```

ضمیمه ب ۰ برنامه کامپیوتری برای الگوریتم ۲

برای توضیح الگوریتم به قسمت ۲۰۳ فصل سوم رجوع کنید.

```
      PROGRAM BL1S
      PARAMETER (N=1000)
      DIMENSION Y(N),X2(N),Z(N),W(N),L(N)
      DO 10 I=1,N
10    READ(5,20) Y(I),X2(I)
20    FORMAT(2F10.3)
      K1=N/2
      K1R=0
      K1S=0
30    DO 40 I=1,K1-1
      W(I)=X2(I)-X2(K1)
40    Z(I)=(Y(I)-Y(K1))/W(I)
      W(K1)=0.
      Z(K1)=0.
      DO 50 I=K1+1,N
      W(I)=X2(I)-X2(K1)
50    Z(I)=(Y(I)-Y(K1))/W(I)
      ITER=ITER+1
      LM=LWMED(N,Z,W,L)
      K1S=K1R
      K1R=K1
      IF (LM.EQ.K1S) GOTO 60
      K1=LM
      GOTO 30
60    B2=Z(LM)
      B1=Y(K1)-B2*X2(K1)
      PRINT 70,B1,B2
70    FORMAT(1X,'B1=',F13.5,3X,'B2=',F13.5)
      STOP
      END
```

برای توضیح الگوریتم به قسمت های ۲۰۴ و ۳۰۴ فصل سوم رجوع کنید.

```
      PROGRAM BL1
      PARAMETER (N=1000,M=5,M1=M-1,M2=M-2)
      DIMENSION Y(N),X(N,M1),XSK(M1),YW(N),XKW(N)
      DIMENSION W(N),YS(N),XS(N,M1),B(M),XW(N,2:M1)
      DIMENSION L(N),KK(M1),YSOL(M1),XSOL(M1,M1)
      COMMON /C1/I1,I2
      DO 10 I=1,N
   10 READ(5,20) Y(I),(X(I,J),J=1,M1)
   20 FORMAT(10F10.3)
      ITER=0
      KR=0
      MM=1
      DO 30 J=1,M1
   30 KK(J)=J*N/M
   40 DO 50 I=1,N
      YS(I)=Y(I)
      DO 50 J=1,M1
   50 XS(I,J)=X(I,J)
      GO TO 80
   60 DO 70 I=1,N
      W(I)=XKW(I)
      YS(I)=YW(I)
      DO 70 J=MM,M1
   70 XS(I,J)=XW(I,J)
   80 JJ=MM
   90 K=KK(JJ)
      YSK=YS(K)
      DO 100 J=JJ,M1
  100 XSK(J)=XS(K,J)
      I1=1
      I2=K-1
  110 DO 120 J=JJ,M1
  120 CALL COL1(XSK(J),XS(1,J))
      CALL COL2(YSK,JJ,W,YS,XS(1,JJ))
      IF(I2.EQ.N) GO TO 130
      I1=K+1
      I2=N
      GO TO 110
  130 W(K)=0.
      IF (JJ.EQ.M1) GO TO 190
      I1=1
      I2=K-1
  140 DO 150 J=JJ+1,M1
  150 CALL COL3(XS(1,J),XS(1,JJ))
      IF(I2.EQ.N) GO TO 160
      I1=K+1
      I2=N
      GO TO 140
  160 IF(JJ.NE.MM) GO TO 180
      DO 170 I=1,N
      XKW(I)=W(I)
      YW(I)=YS(I)
      DO 170 J=JJ+1,M1
  170 XW(I,J)=XS(I,J)
  180 JJ=JJ+1
      GO TO 90
  190 YS(K)=0.
      ITER=ITER+1
      LM=LWMED(N,YS,W,L)
      IF(LM.EQ.KR) GO TO 220
      IOPT=0
  200 IF(MM.EQ.M1) GO TO 210
      MM=MM+1
      KR=KK(MM)
      KK(MM)=LM
      GO TO 60
  210 MM=1
      KR=KK(MM)
      KK(MM)=LM
      GO TO 40
  220 IOPT=IOPT+1
```

```
      IF (IOPT.NE.M1) GO TO 200
      B(M)=YS(LM)
      DO 230 I=1,M1
      YSOL(I)=Y(KK(I))
      DO 230 J=1,M1
230   XSOL(I,J)=X(KK(I),J)
      JJ=1
240   YSK=YSOL(JJ)
      DO 250 J=JJ,M1
250   XSK(J)=XSOL(JJ,J)
      DO 270 I=JJ,M1
      IF(I.EQ.JJ) GO TO 270
      YSOL(I)=YSOL(I)-YSK
      DO 260 J=JJ,M1
260   XSOL(I,J)=XSOL(I,J)-XSK(J)
      YSOL(I)=YSOL(I)/XSOL(I,JJ)
270   CONTINUE
      DO 290 I=JJ,M1
      IF(I.EQ.JJ) GO TO 290
      DO 280 J=JJ+1,M1
280   XSOL(I,J)=XSOL(I,J)/XSOL(I,JJ)
290   CONTINUE
      IF (JJ.EQ.M2) GO TO 300
      JJ=JJ+1
      GO TO 240
300   DO 320 I=1,M2
      K=M-I
      S=YSOL(K)
      DO 310 J=K,M1
310   S=S-B(J+1)*XSOL(K,J)
320   B(K)=S
      S=Y(KK(1))
      DO 330 J=1,M1
330   S=S-B(J+1)*X(KK(1),J)
      B(1)=S
      PRINT 340,(B(J),J=1,M)
340   FORMAT(1X,F13.5)
      PRINT 350,(KK(J),J=1,M1),LM,ITER
350   FORMAT(1X,I13)
      STOP
      END
```

```
      SUBROUTINE COL1(V1,V2)
      DIMENSION V2(1)
      COMMON /C1/I1,I2
      DO 1 I=I1,I2
  1   V2(I)=V2(I)-V1
      RETURN
      END
```

```
      SUBROUTINE COL2(YSK,JJ,V1,YS,V2)
      DIMENSION V1(1),V2(1),YS(1)
      COMMON /C1/I1,I2
      IF (JJ.NE.1) GO TO 2
      DO 1 I=I1,I2
      V1(I)=V2(I)
  1   YS(I)=(YS(I)-YSK)/V2(I)
      RETURN
  2   DO 3 I=I1,I2
      V1(I)=V1(I)*V2(I)
  3   YS(I)=(YS(I)-YSK)/V2(I)
      RETURN
      END
```

```
      SUBROUTINE COL3(V1,V2)
      DIMENSION V1(1),V2(1)
      COMMON /C1/I1,I2
      DO 1 I=I1,I2
  1   V1(I)=V1(I)/V2(I)
      RETURN
      END
```

"کتابشناسی"

N.N. Abdelmalek (1971) Linear approximation for a discrete point set and L_1 solutions of overdetermined linear equations. J. ACM, 18, 41-47.

N.N. Abdelmalek (1974) On the discrete linear L_1 approximation and L_1 solutions of overdetermined linear equations. J. of Approx. Theory, 11, 38-53.

N.N. Abdelmalek (1975a) An efficient method for the discrete L_1 approximation problem. Math. Comput., 29, 844-850.

N.N. Abdelmalek (1975b) Chebyshev solution of overdetermined system of linear equations. BIT, 15, 117-129.

N.N. Abdelmalek (1976) A computer program for the Chebyshev solution of overdetermined system of linear equations, Inter. J. Numer. Meth. in Engin. 10, 1197-1202.

N.N. Abdelmalek (1977a) Computing the strict Chebyshev solution of overdetermined linear equations, Math. Comp., 31, 974-983.

N.N. Abdelmalek (1977b) The discrete linear restricted Chebyshev approximation. BIT, 17, 249-261.

N.N. Abdelmalek (1980a) L_1 solution of overdetermined systems of linear equations. ACM Trans. Math. Soft., 6, 220-227.

N.N. Abdelmalek (1980b) A Fortran subroutine for the L_1 solution of overdetermined systems of linear equations. ACM Trans. Math. Soft., 6, 228-30.

S. Abou-Jaoude (1976a) Sur une condition necessaire et suffisante de L_1-convergence pres que complete de l'estimateur de la partition fixe pour une densite. C. R. Acad. Sci. Paris, Ser. A 283, 1107-1110.

S. Abou-Jaoude (1976b) Sur la convergence L_1 at L_∞ de l'estimateur de la partition aleatoire pour une densite. Ann. Inst. Henri Poincare, 12, 299-317.

S. Abou-Jaoude (1976c) Conditions necessaires et suffisantes de convergence L_1 en probabilite de l'histogramme pour une densite. Ann. Inst. Henri Poincare, 12, 213-231.

J.F. Affleck-Graves, A.H. Money, K. Carter (   ) An evaluation of an alternative methods of estimating the beta coefficient in the market model. University of CapeTown, South Africa.

T. Amemiya (1982) Two stage least absolute deviations estimators. Econometrica, 50,3, 689-711.

T.W. Anderson (1962) Least squares and best unbiased estimates. Annals of Math. Stat., 33, 266-272.

D.W. Anderson (1965) Linear programming time estimating equations. J. of Indus. Engin. 16, 136-138

D.H. Anderson (1975) Linear programming and the calculation of maximum norm approximations. Ph.D. thesis, Australian National University.

D.H. Anderson, M.R. Osborne (1976) Discrete linear approximation problems in polyhedral norms. Numer. Math. 26, 179-189.

D.H. Anderson, M.R. Osborne (1977a) Discrete nonlinear approximation problems in polyhedral norms, a Levenberg-like algorithm. Numer. Math. 28, 157-170.

D.H. Anderson, M.R. Osborne (1977b) Discrete nonlinear approximation in polyhedral norms. Numer. Math., 28, 143-156.

D. Anderson, W.L. Steiger (1982) A comparison of methods for discrete L_1 curve-fitting. Tech. Rep. DCS-TR-96. Dept. of Comp. Sci., Hill center for the Math. Sci. Busch campus, New Brunswick, N.J.

F.J. Anscombe (1976) Topics in the investigation of linear relations fitted by the method of least square (with discussion), J. of Royal Stat. Soc. B series, 1-52.

J. Antoch, A. Bartkowiak, J. Pekalska (1986) Computing L_1 norm, α-trimmed and α-winsorized regression on the ODRA 1305 computer. Rep. N-159, Institute of Computer Science, Wroclaw university, Poland.

J. Antoch (1987) Variable selection in linear model based on trimmed least squares estimator. In Y. Dodge (ed.) Statistical data analysis based on the L_1 norm and related methods. North-Holland. 231-246.

H. Anton, C.S. Duris (1973) On an algorithm for best approximate solutions to **Av=b** in normed linear spaces. J. Approx. Theory 8, 133-141.

M. Aoki (1965) Successive generation of Chebyshev approximate solution. J. Basic Engin., 87, 17-22.

G. Appa, C. Smith (1973) On L_1 and Chebyshev estimation. Math. Prog. 5, 73-87.

R.D. Armstrong, J.J. Elam, J.W. Hultz (1977) Obtaining least absolute value estimates for a two-way classification model. Commun. Statist., B6, 365-81.

R.D. Armstrong, E.L. Frome (1976a) A comparison of two algorithms for absolute deviation curve fitting. JASA, 71, 328-330.

R.D. Armstrong, E.L. Frome (1976b) The calculation of least absolute value estimates for two-way tables. Proceeding of the statistical computing section, ASA, Washington D.C., 101-106.

R.D. Armstrong, E.L. Frome (1977) A special purpose linear programming algorithm for obtaining least absolute value estimates in a linear model with dummy variables. Commun. Stat., B6, 383-98.

R.D. Armstrong, E.L. Frome (1979) Least-absolute-value estimators for one-way and two-way tables. Naval Res. Log. Quart., 26, 79-96.

R.D. Armstrong, E.L. Frome, D.S. Kung (1979) A revised simplex algorithm for the absolute deviation curve fitting problem. Commun. Stat. B8, 175-190.

R.D. Armstrong, E.L. Frome, M.G. Sklar (1980) Linear programming in exploratory data analysis. J. of Educ. Stat., 5, 293-307.

R.D. Armstrong, J. Godfrey (1979) Two linear programming algorithms for the discrete L_1 problem. Math. Comput., 33, 289-300.

R.D. Armstrong, J.W. Hultz (1977) An algorithm for a restricted discrete approximation problem in L_1 norm. SIAM J. Numer. Anal., 14, 555-565.

R.D. Armstrong, D.S. Kung (1978) AS132: Least absolute value estimates for a simple linear regression problem. Appl. Stat., 27, 363-366.

R.D. Armstrong, D.S. Kung (1979) AS135: Mini-max estimates for a linear multiple regression problem. Appl. Stat., 93-100.

R.D. Armstrong, D.S. Kung (1980) An algorithm for a least absolute value regression problem with bounds on the parameters. Appl. Math. Comput. 7, 267-279.

R.D. Armstrong, D.S. Kung (1982a) An algorithm to select the best subset for a least absolute value regression problem. TIMS studies in the management sciences, 19, 67-80.

R.D. Armstrong, D.S. Kung (1982b) A dual algorithm to solve linear least absolute value problems. J. Oper. Res. Soc., 33, 931-936.

K.J. Arrow, M. Hoffenberg (1959) A time series analysis of interindustry demands. North-Holland, Amsterdam.

T.S. Arthanari, Y. Dodge (1981) Mathematical programming in statistics. John Wiley, Interscience division, New York.

W.C. Ashar, T.D. Wallace (1963) A sampling of minimum absolute deviations estimators. Oper. Res., 11, 747-752.

P. Assouad (1977) Un espace hypermetric non plongeable dans un espace L_1. C. R. Acad. Sc. Paris, T. 285, Serie A, 361-363.

S. Baboolal, G.A. Watson (1981) Computational experience with an algorithm for discrete L_1 approximation. Computing, 27, 245-252.

S.C. Banks, H.L. Taylor (1980) A modification to the discrete L_1 linear approximation algorithm of Barrodale and Roberts. SIAM J. on Scientific and Stat. Comput. 1, 187-190.

G.D.I. Barr, J.F. Affleck-Graves, A.H. Money, M.L. Hart (1980a) Performance of a generalized algorithm for L_p-norm regression estimates. Tech. Rep. no. ALS-4, Aug., University of CapeTown, Dept. of Math. Stat. South Africa.

G.D.I. Barr, J.F. Affleck-Graves, A.H. Money, M.L. Hart (1980b) L_p norm estimation and the choice of p, a practical approach. University of CapeTown, Dept. of Math. Stat., Tech. Rep. no. ALS-3, july.

G.D.I. Barr, A.H. Money, J.F. Affleck-Graves, M.L. Hart (1980) L_p-norm estimation of a symmetric distribution. University of CapeTown, Dept. of Math. Stat., Tech. Rep. no. ALS-5, Oct..

G.D.I. Barr, A.H. Money, J.F. Affleck-Graves, M.L. Hart (1981a) Estimation of location for skewed data sets: a comparative study utilizing the data set published by Stigler. University of CapeTown, Dept. of Math. Stat., Tech. Rep. no. ALS-7, may.

G.D.I. Barr, A.H. Money, J.F. Affleck-Graves, M.L. Hart (1981b) Estimation of location for skewed data sets. University of CapeTown, Dept. of Math. Stat., Tech. Rep. no. ALS-6, April.

I. Barrodale (1967) Approximation in the L_1 and L_∞ norms by linear programming. Ph.D. thesis, University of Liverpool, Liverpool, England.

I. Barrodale (1968) L_1 approximation and analysis of data. Appl. Stat. 17,51-57.

I. Barrodale (1970) On computing best L_1 approximations. In A. Talbot, Approximation theory Academic Press, New York, 205-215.

I. Barrodale, C. Phillips (1975) Solution of an over-determined system of linear equations in Chebyshev norm. ACM Trans. on Math. Soft. 264-70.

I. Barrodale, M.J.D. Powell, F.D.K. Roberts (1972) The differential correction algorithm for rational L_∞ approximation. SIAM J. Num. Anal., 9, 493-503.

I. Barrodale, F.D.K. Roberts (1970) Application of mathematical programming to L_p approximation. In J.B. Rosen, O.L. Mangasarian, K. Ritter, Nonlinear programming Academic Press, New York, 447-64.

I. Barrodale, F.D.K. Roberts (1973) An improved algorithm for discrete L_1 linear approximation. SIAM J. Numer. Anal., 10, 839-848.

I. Barrodale, F.D.K. Roberts (1974) Algorithm 478: Solution of an overdetermined system of equations in the L_1 norm. Commun. ACM, 17, 319-320.

I. Barrodale and F.D.K. Roberts (1977) Algorithms for restricted least absolute value estimation. Commun. Stat. B6(4), 353-363.

I. Barrodale and F.D.K. Roberts (1978) An efficient algorithm for discrete L_1 linear approximation with linear constraints. SIAM J. Numer. Anal., 15, 603-611.

I. Barrodale and F.D.K. Roberts, C.R. Hunt (1970) Computing best L_p approximations by functions nonlinear in one parameter. Comp. J., 13, 382-386.

I. Barrodale, C. Phillips (1974) An improved algorithm for discrete Chebyshev linear approximation. Proc. 4th Manitoba conf. on numer. math. University of Manitoba, Winnipeg, Manitoba, 177-190.

I. Barrodale, A. Young (1966) Algorithms for best L_1 and L_∞ linear approximations on a discrete set. Numer. Math., 8, 295-306.

R.H. Bartels, A.R. Conn (1977) LAV regression: A special case of piecewise linear minimization. Commun. Stat., B6, 329-340.

R.H. Bartels, A.R. Conn (1980a) Linearly constrained discrete L_1 problems. ACM Trans. Math. Soft. 6, 594-608.

R.H. Bartels, A.R. Conn (1980b) Algorithm 563, A program for linearly constrained discrete L_1 problems. ACM trans. Math. Soft., 6, 609-614.

R.H. Bartels, A.R. Conn (1982) An approach to nonlinear L_1 data fitting. In J.P. Hennart (ed.), Numerical analysis, Cocoyoc, Springer Verlag, 45-58.

R.H. Bartels, A.R. Conn, C. Charalambous (1978) On Cline's direct method for solving overdetermined linear system in the L_∞ sense, SIAM J. Numer. Anal., 15, 255-270.

R.H. Bartels, A.R. Conn, Y. Li (1987) Primal methods are better than dual methods for solving overdetermined linear systems in the L_∞ sense? Res. Rep. CS-87-12, University of Waterloo, Comp. Sci. Dept.

R.H. Bartels, A.R. Conn, J. Sinclair (1976) The L_1 solution to an overdetermined linear system. Proc. 9th Ann. Symp. Interface Statist. In C.D.C. Hoaglin (ed.) Boston, Prindle, Weber and Schmidt Inc., 120-7.

R.H. Bartels, A.R. Conn, J. Sinclair (1978) Minimization technique for piecewise differentiable functions: The L_1 solution to an overdetermined linear system. SIAM J. Numer. Anal. 15, 224-241.

R.H. Bartels, G.H. Golub (1968a) Algorithm 328: Chebyshev solution to an overdetermined linear system. Commun. ACM 11, 428-430.

R.H. Bartels, G.H. Golub (1968b) Stable numerical methods for obtaining the Chebyshev solution to an overdetermined system of equations, Commun. ACM, 11, 401-406.

R.H. Bartels, G.H. Golub (1969) The simplex method of linear programming using LU decomposition. Commun. ACM, 12, 266-268.

G.W. Bassett (1973) Some properties of the least absolute error estimator. Ph.D. thesis. Dept. of Econ., University of Michigan.

G.W. Bassett, Jr. (1987) A p-subset property of L_1 and regression quantile estimates. Work. Pap., Dept. of Economics, University of Illinois-Chicago.

G.W. Bassett, Jr. (1988a) A p-subset property of L_1 and regression quantile estimates. CSDA, 6(3), 297-304.

G.W. Bassett, Jr. (1988b) A property of the observations fit by the extreme regression quantiles. CSDA, 6(4), 353-360.

G.W. Bassett, Jr., R. Koenker (1978) Asymptotic theory of least absolute error regression. JASA, Sep., 73, no. 363, 618-22.

G.W. Bassett, Jr., R. Koenker (1982) An empirical quantile function for linear models with iid errors. JASA, 77, no. 378, 407-415.

J. Bejar (1956) Regression en mediana y la programacion lineal, Trabajos de Estadistica 7, 141-58.

J. Bejar (1957) Calculo practico de la regression en mediana Trabajos de Estadistica, 8, 157-173.

C.M. Bender, S.A. Orszag (1978) Advanced mathematical methods for scientists and engineers, McGraw-Hill.

B. Bidabad (1980) The new international scholastic order, an essay in political economy of science. Shiraz University, Shiraz, Iran.

B. Bidabad (1981) Estimation of the Engel's curves for Iran (urban & rural). M.S. thesis. Dept. of Econ., Shiraz University, Shiraz, Iran.

B. Bidabad (1983a) Analysis of taxonomy and its application in classification of regions and creating development index for regional planning. Plan and Budget Ministry, Centeral Province Plan and Budget Organization. Arak, Iran.

B. Bidabad (1983b) The discorrelation of provincial development budget allocation for 1984 and undevelopment index. Centeral Province Plan and Budget Organization, Plan and Budget Ministry, Arak, Iran.

B. Bidabad (1983c) Principal components analysis in creating development index for regional planning. Regional Planning Bureau, Plan and Budget Ministry, Tehran Iran.

B. Bidabad (1984a) Goal programming, optimal decision process with different priorities and its computer program. Plan and Budget Ministry, Regional Planning Bureau, Tehran, Iran.

B. Bidabad (1984b) Determining industrial location for 1992 by using goal programming method. Plan and Budget Ministry, Regional Planning Bureau, Tehran, Iran.

B. Bidabad (1984c) Zero-one programming, optimal decision making with bivalent variables and its computer program. Plan and Budget ministry, Regional Planning Bureau, Tehran, Iran.

B. Bidabad (1985) Model of Iran's economic growth alternatives, 1986-2001 prospect. Plan and Budget Ministry, Regional Planning Bureau, Tehran, Iran.

B. Bidabad (1987a) A proposed algorithm for least absolute error estimation. Submitted to the First International Conference on Statistical Data Analysis Based on the L_1 norm and Related Methods, Neuchatel, Switzerland.

B. Bidabad (1987b) A proposed algorithm for least absolute error estimation, part II. Submitted to the First International Conference on Statistical Data Analysis Based on the L_1 norm and Related Methods, Neuchatel, Switzerland.

B. Bidabad (1988a) A proposed algorithm for least absolute error estimation. Proceedings of the Third Seminar of Mathematical Analysis. Shiraz University, 24-34, Shiraz, Iran.

B. Bidabad (1988b) A proposed algorithm for least absolute error estimation, part II. Proceedings of the Third Seminar of Mathematical Analysis, Shiraz University, 35-50, Shiraz, Iran.

B. Bidabad (1989) A mathematical model for simultaneous computation of preferred prices and number of pieces of lands in non-beneficial land preparation projects. Economics and Management quarterly journal of the Islamic Azad University. Vol.2, 38-50.

B. Bidabad ( ) Dynamic economic models. vol. I, Difference equations. Unpublished.

B. Bidabad, B. Bidabad (1989a) Functional form for estimating the Lorenz curve. Submitted to the Australasian Meeting of Econometric Society, Canberra, Australia.

B. Bidabad, B. Bidabad (1989b) Functional form for estimating the Lorenz curve. To be appeared in Economics and Management quarterly journal of the Islamic Azad University.

B. Bidabad, B. Bidabad (1989c) Complex probability and Markov stochastic processes. Submitted to the 47th session of ISI, sep., Paris, France.

B. Bidabad, B. Bidabad (1989d) Complex probability and Markov stochastic process. To be appeared in Economics and Management quarterly journal of the Islamic Azad University.

B. Bidabad, M. Sabetghadam, S. Davar, G.R. Sadeghi (1988) Dynamic model of population, mathematical model for analysis and forecasting Iran population changes. Regional Planning Bureau, Plan and Budget Ministry,. Tehran, Iran.

R. Blattberg, T. Sargent (1971) Regression with non-Gaussian stable disturbances: some sampling results, Econometrica, May, 501-510.

P. Bloomfield, W. Steiger (1980) Least absolute deviations curve fitting. SIAM J. Sci. Statist. Comput. 1, 290-301.

P. Bloomfield, W. Steiger (1983) Least absolute deviations: theory, applications and algorithms. Birkhauser, Boston.

P.T. Boggs (1974) A new algorithm for Chebyshev solution of overdetermined linear system. Math. Comp., 28, 203-218.

M.S. Borowsky (1976) Algorithms for solving the dual problem for **Av=b** with varying norms. J. Approx. Theory, 17, 107-114.

R.J. Boscovich (1757) De litteraria expeditione per pontificiam ditionem, et synopsis amplioris operis..., 'Bononiensi' scientiarum et artum instituto atque academia commetarii, vol.4, 353-396. Reprinted with a Serbo-Croatian translation by N. Cubranic, Institute of higher geodesy, University of Zagreb 1961.

R.J. Boscovich (1760) De recentissimis graduum dimensionibus et figura, ac magnitudine terrae inde derivanda. Philosophiae recentioris, a benedicto stay in Romano archigynasis publico eloquentare professore, vesibus traditae, libri X, cum adnotianibus et supplementas P. Rugerii Joseph Boscovich, S.J., 2, 406-426.

A.L. Bowley (1902) Methods of representing the statistics of wages and other groups not fulfilling the normal law of error, II: applications to wage statistics and other groups. J. of the Roy. Stat. Soc., 65, 331-54.

A.L. Bowley (1928) F.Y. Edgeworth's contributions to mathematical statistics. London, Roy. Stat. Soc..

G.E.P. Box, G.C. Tiao (1962) A further look at robustness vi Bayes' theorem. Biometrika, 419-432.

D. Bradu (1987a) L_1 fit, median polish and conjugate gradients. CSIR Tech. Rep. TWISK 509, National Res. Inst. for Math Sci. CSIR, Pretoria.

D. Bradu (1987b) An ϵ-median polish algorithm. CSDA, 5, 327-

336.

M. Brannigan, S.A. Gustafson (1987) H-sets in convex programming and constrained approximation. Work. Pap., Rogaland University, Stavanger.

H.D. Brecht (1976) Regression methodology with observation errors in the explanatory variables. Decis. Sci. 7, 57-65.

J.J. Brennan, L.M. Seiford (1987) Linear programming and L_1 approximation using the method of vanishing Jacobians. CSDA, 5, 263-276.

B.M. Brown (1980) Median estimates in a simple linear regression. Australian J. of Stat., 22, 154-165.

B.M. Brown, T.P. Hettmansperger (1987) Invariant tests in bivariate models and the L_1-criterion. In Y. Dodge (ed.) Statistical data analysis based on the L_1 norm and related methods, North-Holland, 333-344.

C. Bruen (1938) Methods for the combination of observations modal points or most lesser-deviations, mean loci or least squares, and mid point of least range or least greatest-deviation. Metron 13, 61-140.

J.J. Buckley, A.H. Kvanli (1981) The MAD estimator: when is it non-linear? Applications to two-way design models. Commun. Stat. A10, 2581-2590.

Burgoyne (1965) Polynomial approximation by Edgeworth's method. University of London.

S. Busovaca (1985) Handling degeneracy in a nonlinear L_1 algorithm. Ph.D. thesis, University of Waterloo, Waterloo, Ontario.

V.A. Cabot, R.L. Francis, M.A. Stary (1970) A network flow solution to a rectilinear distance facility location problem. AIIE trans., vol. 2.

P.H. Calamai, A.R. Conn (1987) A projected Newton method for L_p norm location problems. Math. Prog. 38, 75-109.

G. Le Calve (1987) L_1-embeddings of data structure (I,D). In Y. Dodge (ed.) Statistical data analysis based on the L_1 norm and related methods. 195-202., North-Holland.

J. Chambers (1971) Algorithm 410: partial sorting. Commun. ACM, 14, 357-358.

J.M. Chambers (1977) Computational methods for data analysis Wiley, New York.

D.G. Champernowne (1953) A model of income distribution. Economic J., 63, 318-351.

A. Charnes, W.W. Cooper, R.O. Ferguson (1955) Optimal estimation of executive compensation by linear programming. Manag. Sci. 1, 138-151.

C. Charalambous, A.R. Conn (1978) An efficient method to solve the minimax problem directly. SIAM J. Numer. Anal., 15, no.1, 162-187.

C. Charalambous (1979) On conditions for optimality of nonlinear L_1 problem. Math. Prog. 17, 123-135.

E.W. Cheney (1966) Introduction to approximation theory McGraw-Hill, New York.

W. Cheney, A.A. Goldstein (1958) Note on a paper by Zuhovickii concerning the Tchebysheff problem for linear equations. SIAM J. Numer. Anal. 6, 233-239.

J.A. Chisman (1966) Using linear programming to determine time standards, J. Ind. Engin. 17, 189-191.

Su Chun (1987) Some results on the L_p-convergence ($p \geq 1$) of U-statistics. CSDA, 5, 321-326.

D.I. Clarke (1981) Finite algorithms for linear optimization problems. Ph.D. thesis, Australian National University, Canberra.

F.H. Clarke (1983) Optimization and nonsmooth analysis. Wiley, New York

J.F. Claerbout, F. Muir (1973) Robust modeling with erratic data. Geophysics 38, 826-844.

A.K. Cline (1970) Uniform approximation as a limit of approximations. Ph.D. thesis, University of Michigan, Ann, Arbor, Michigan.

A.K. Cline (1972) Rate of convergence of Lawson's algorithm, Math. of Comp., 26, 167-176.

A.K. Cline (1976) A descent method for the uniform solution of overdetermined system of linear equations. SIAM J. Numer. Anal., 13, 293-309.

K.O. Cogger (1979) Time-series analysis and forecasting with an absolute error criterion. TIMS Studies in Manag. Sci., S. Markidakis, S.C. Wheelwright (eds.) North-Holland, 189-201, Amsterdam.

T.F. Coleman (1978) A note on 'New algorithms for constrained minimax optimization', Math. Prog., 15, 239-242.

A.R. Conn (1975) Optimization of microwave networks. IEEE Tran. on Microwave Theory and Techniques, Oct., 834-838.

A.R. Conn (1976) linear programming via a nondifferentiable penalty function. SIAM J. Numer. Anal., 13, 145-154.

A.R. Conn (1984) Nonlinear programming, exact penalty functions and projection techniques for non-smooth functions. Tech. Rep. no. CS-84-26. Dept. of Comp. Sci., University of Waterloo, Ontario, Canada.

A.R. Conn, N.I.M. Gould (1987) An exact penalty function for semi-infinite programming. Math. Prog., 37, 19-40.

B. Cornell, J.K. Dietrich (1978) Mean-absolute-deviation versus least-squares regression estimation of beta coefficients. J. of Financial and Quantitative analysis 13, 123-131.

F. Critchley (1980) Optimal norm characterizations of multidimensional scaling methods and some related data analysis problem. In E. Diday et al (eds.) Data analysis and informatics, North-Holland, 209-229.

D.C. Crocker (1969) Linear programming technique in regression analysis, the hidden danger. A.I.E.E. Trans., 1, 112-126.

J.S. Cramer (1973) Empirical econometrics. North-Holland, Amsterdam.

M. Csorgo, L. Horvath (1987) Asymptotics for L_p-norms of naive estimators of densities. Tech. Rep. series of Lab. Res. Stat. Prob. no. 96, Carleton University, Ottawa, Canada.

M. Csorgo, L. Horvath (1988) Asymptotics for L_p-norms of kernel estimators of densities. CSDA, 6(3), 241-250.

M. Csorgo, E. Gombay, L. Horvath (1987) Asymptotics for L_p-norms of kernel estimators of density under random censorship. Tech. Rep. series of Lab. Res. Stat. Prob., Carleton University, Ottawa, Canada.

M. Davies (1976) Linear approximation using the criterion of least total deviations. J. Roy. Stat. Soc. B29, 101-109.

J. Descloux (1963) Approximation in L_p and Chebyshev approximations. SIAM J. 11, 1017-1026.

L. Devroye (1983) The equivalence of weak, strong and complete convergence in L_1 for kernel density estimates. Annals of Stat., 11, 896-904.

L. Devroye (1985) A note on the L_1 consistency of variable kernel estimates. Ann. Stat., 13, 1041-1049.

L. Devroye, L. Gyorfi (1985) Non parametric density estimation, the L_1 view, Wiley, New York.

L. Devroye, T.J. Wagner (1979) The L_1 convergence of kernel density estimates. Ann. Stat., 7, 1136-1139.

L. Devroye, T.J. Wagner (1980) On the L_1 convergence of kernel estimators of regression functions with applications in discrimination. Z. Wahrsch, verw. Gebiete, 51, 15-25.

T.E. Dielman (1984) Least absolute value estimation in regression models: An annotated bibliography. Comm. Stat. 13, 513-41.

T. Dielman, R. Pfaffenberger (1982) LAV (Least Absolute Value) estimation in linear regression: A review, TIMS studies in the Manag. Sci., 19, 31-52.

T. Dielman, R. Pfaffenberger (1984) Computational algorithms for calculating least absolute value and Chebyshev estimates for multiple regression. Amer. J. Math. Manag. Sci., 4, 169-197.

J. Dieudonne (1963) Foundations of modern analysis. Academic Press, New York.

J.J. Dinkel, R.C. Pfaffenberger (1981) Constrained L_1 estimation via geometric programming. European J. Oper. Res. 7, 299-305.

P.J. Dhrymes (1978) Mathematics for econometrics. Springer-Verlag, New York.

Y. Dodge (1984) Robust estimation of regression coefficients by minimizing a convex combination of least squares and least absolute deviations. Comp. Stat. Quarterly, 1, 139-153.

Y. Dodge (1987) An introduction to statistical data analysis L_1-norm based. In Y. Dodge (ed.) Statistical data analysis based on the L_1 norm and related methods. North-Holland. Reprinted in CSDA, 5, 239-254.

Y. Dodge, J. Jureckova (1987) Adaptive combination of least squares and least absolute deviations estimators. In Y. Dodge (ed.) Statistical data analysis based on the L1-norm and related methods. 275-287, North-Holland.

Y. Dodge, J. Jureckova (1988) Adaptive combination of M-estimator and L_1-estimator. In Y. Dodge, V.V. Fedorov, H.P. Wynn (eds.) Optimal design and analysis of experiments, 167-176, North-Holland.

W. Domschke, A. Drext (1984) An international bibliography for location- and layout-planning. Instut fuer Unternehmensforschung und Informatik, Hochschule der Eundeswehr, postach 700822 D-2000 Hamburg 70.

E.L. Dougherty, S.T. Smith (1966) The use of linear programming to filter digitized map data. Geophysics, 31, 253-259.

Z. Drezner, G.O. Wesolowsky (1978) A trajectory method for the optimization of multi-facility location problem with L_p distances. Manag. Sci., 24, no. 14, 1507-1514.

A.F. Dufton (1928) Correlation. Nature, 121, 866.

J. Dupacova (1987a) Asymptotic properties of restricted L_1-

estimates of regression. Feb., WP-87-18, Work. Pap. Inter. Inst. for Applied Sys. Analysis, A-2361 Laxenburg, Austria.

J. Dupacova (1987b) Asymptotic properties of restricted L_1-estimates of regression. In Y. Dodge (ed.) statistical data analysis based on the L_1-norm and related methods. 263-274, North-Holland.

C.S. Duris, V.P. Sreedharan (1968) Chebyshev and L_1 solutions of linear equations using least squares solutions. SIAM J. Num. Anal., 5, 491-505.

C.S. Duris, M.G. Temple (1973) A finite step algorithm for determining the "strict" Chebyshev solution to **Ax=b**. SIAM J. Num. Anal. 10, 690-699.

R. Dutter (1977) Numerical solution of robust regression problems, computational aspects, a comparison. J. of Stat. Computation and Simulation, 5, 207-238.

R. Dutter (1987) A fortran program for robust and bounded influence regression. In Y. Dodge (ed.) Statistical data analysis based on the L_1 norm and related methods. 139-144. North-Holland.

F.Y. Edgeworth (1883) The method of least squares. Philosophical Magazine, 16, 360-375.

F.Y. Edgeworth (1887a) On observations relating to several quantities. Hermathena, 6, 279-285.

F.Y. Edgeworth (1887b) A new method of reducing observations relating to several quantities. Philosophical Magazine, 24, 222-223.

F.Y. Edgeworth (1888) On a new method of reducing observation relating to several quantities. Philosophical Magazine, 25, 184-191.

F.Y. Edgeworth (1902) Method of representing statistics of wage and other groups not fulfilling the normal law of error, I: mathematical considerations. J. Roy. Stat. Soc., 65, 325-331.

F.Y. Edgeworth (1923) On the use of medians for reducing observations relating to several quantities. Philosophical Magazine, 6th series, 46, 1074-1088.

C. Eisenhart (1961) Boscovich and the combination of observations. Ch. 9 of Whyte (1961, 200-212) reprinted in Kendall and Plackett (1977) studies in the history of statistics and probability, vol.II, Charles Griffin and Co. Ltd., High Wycombe 88-100.

H. Ekblom (1973a) A note on nonlinear median estimators. JASA, 68, 431-2.

H. Ekblom (1973b) Calculation of linear best L_p-approximations. BIT 13, 292-300.

H. Ekblom (1974) L_p methods for robust regression. BIT 14, 22-32.

H. Ekblom (1987) The L_1 estimate as limiting case of an L_p- or Huber estimate. In Y. Dodge (ed.) Statistical data analysis based on the L_1 norm and related methods. North-Holland. 109-116.

H. Ekblom, S. Henriksson (1969) Lp-criteria for the estimation of location parameters. SIAM J. Appl. Math., 17, 1130-1141.

R.A. El-Attar, M. Vidyasagar, S.R.K. Dutta (1976) Optimality conditions for L_1-norm minimization Proc. 19th midwest symposium on circuits and systems, 272-275.

R.A. El-Attar, M. Vidyasagar, S.R.K. Dutta (1979) An algorithm for L_1-norm minimization with application to nonlinear L_1-approximation. SIAM J. Numer. Anal., 16, 70-86.

J.E. Estienne (1926-28) Introduction a une theorie rationnelle des erreurs d'observation. Revue d'artillerie 97(1926), 421-441; 98(1928), 542-562; 100(1927), 471-487.

R.C. Fair (1974) On the robust estimation of econometric models. Ann. Econ. Soc. Measurement, 3, 667-77.

E. Fama (1965) The behavior of stock market prices. J. Bus. 38, 34-105.

E.F. Fama, R. Roll (1971) Parameter estimates for symmetric stable distributions. JASA 66, 331-338.

R.W. Farebrother (1985) Unbiased L_1 and L_∞ estimation. Commun. Stat., A, 14,1941-1962.

R.W. Farebrother (1987a) The theory of committee decisions and the double median method. CSDA, 5, 437-442.

R.W. Farebrother (1987b) The historical development of the L_1 and L_∞ estimation procedures. In Y. Dodge (ed.) Statistical data analysis based on the L_1 norm and related methods. North-Holland. 37-64.

R.W. Farebrother (1987c) A simple recursive procedure for the L_1 norm fitting of a straight line. Work. Pap., Dept. of Econometrics and Social Stat. University of Manchester, Manchester, M13 9PL, UK.

R.W. Farebrother (1987d) Mechanical representation of the L_1 and L_2 estimation problems. In Y. Dodge (1987) Statistical data analysis based on the L_1 norm and related methods. North-Holland. 455-464.

R.W. Farebrother (1987e) A remark on AS108: multiple linear regression with minimum sum of absolute errors. Appl. Stat., 36, no. 1, 118.

D.W. Fausett, J.H. Weber (1978) Mass spectral pattern recognition via techniques of mathematical programming. Analytical Chemistry, 50, 722-731.

V.V. Fedorov (1987) Various discrepancy measures in model testing (two competing regression models). In Y. Dodge (ed.) Statistical data analysis based on the L_1 norm and related methods. North-Holland, 357-366.

B. Fichet (1987a) The role played by L_1 in data analysis. In Y. Dodge (ed.), Statistical data analysis based on the L_1 norm and related methods. North-Holland, 185-194.

B. Fichet (1987b) L_p-space in data analysis. First Conference of the International Federation of Classification Societies. Aachen.

W.D. Fisher (1961) A note on curve fitting with minimum deviations by linear programming. JASA, 11, 359-362.

R. Fletcher (1981) Numerical experiments with an exact L_1 penalty function method. In O.L. Mangasarian, R.R. Meyer, S.M. Robinson (eds.) Nonlinear programming 4. Academic Press, New York, 99-129.

R. Fletcher (1984) An L_1 penalty method for nonlinear constraints. Rep. NA/81, Dept. of Math. Sci., University of Dundee, Scotland.

R. Fletcher, J.A. Grant, M.D. Hebden (1971) The calculation of best L_p approximations. Comp. J., 14, 276-279.

R. Fletcher, J.A. Grant, M.D. Hebden (1974a) Linear minimax approximation as the limit of best L_p approximation. SIAM J. Numer. Anal., 11, 123-136.

R. Fletcher, J.A. Grant, M.D. Hebden (1974b) The continuity and differentiability of parameters of best linear L_p approximations. J. Approx. Theory, 10, 69-73.

A.B. Forsythe (1972) Robust estimation of straight line regression coefficients by minimizing p^{th} power deviations. Technometrics, 14, 159-166.

C.R. Forth (1974) Robust estimation techniques for population parameters and regression coefficients. M.S. thesis. Air Force Institute of Technology, Wright-Patterson,AFB, Ohio.

R. Fourer (1985a) A simplex algorithm for piecewise-linear programming I: derivation and proof. Math. Prog., 33, 204-233.

R. Fourer (1985b) A simplex algorithm for piecewise-linear programming II: Finiteness, feasibility and degeneracy. Tech. Rep., 85-03 (revised), Dept. of Ind. Engin. and Manag. Sci., The Tech. Inst., Northwestern University, Evanston, Illinois.

R. Fourer (1986) A simplex algorithm for piecewise-linear programming III: Computational analysis and applications. Tech. Rep., 86-03, Dept. of Ind. Engin. and Manag. Sci., The Tech. Inst., Northwestern University, Evanston, Illinois.

J.B.I. Fourier (1824) Solution d'une question particuliere au calcul des inegalites, second extrait. Histoire de l'academie des sciences pour 1824, 47-55. Reprinted in oeuvres de Fourier, 2. Paris, 1980, Gauthier-Villars, 325-328.

E.L. Frome, R.D. Armstrong (1977) A robust procedure for estimating the trend-cycle component of an economic time series. In D. Hogben (ed.) Proceeding of the tenth symposium on the interface. Gaithersburg: National Bureau of Standards.

E.L. Frome, G.J. Yakatan (1980) Statistical estimation of the pharmokinetic parameters in the one compartment open model. Commun. Stat. B9, 201-222.

A. Gaivoronski (1987) Numerical techniques for finding estimates which minimize the upper bound of absolute deviations. In Y. Dodge (ed.) Statistical data analysis based on the L_1 norm and related methods. North-Holland. 247-262.

G. Galilei (1632) Dialogo dei massimi sistemi.

J.S. Galpin (1986) Robust and bounded influence regression. National Res. Inst. for Math. Sci. WNNR, CSIR, TWISK, Pretoria, South Africa.

J.S. Galpin, D.M. Hawkins (1987) Methods of L_1 estimation of a covariance matrix. CSDA, 5, 305-319.

C.W. Ganger, D. Orr (1972) Infinite variance and research strategy in time series analysis. JASA 67, 275-285.

C.B. Garcia, F.G. Gould (1983) An application of homotopy to solving linear programs. Math. Prog. 27, 263-282.

C.F. Gauss (1809) Theoria motus corporum coelestium. In F. Perthes, I.H. Besser, Sectionbus conicis solem ambientium, Hamburg. Reprinted in his werke, vol. 7, F. Pethes, Gotha 1871. English translation by C.H. Davis, Little, Brown and Co., Boston, 1857. Reprinted by Dover Pub. New York, 1963.

J.E. Gentle (1977) Least absolute value estimation: an introduction. Commun. Stat., B6, 313-28.

J.E. Gentle, T.A. Hansen (1977) Variable selection under L_1, Proceedings of the statistical computing section A.S.A., 228-230.

J.E. Gentle, W.J. Kennedy, V.A. Sposito (1976) Properties of the L_1 estimate space. Proc. Stat. Comp. section A.S.A. 163-164.

J.E. Gentle, W.J. Kennedy, V.A. Sposito (1977) On least absolute values estimation. Comm. Stat., A6, 839-845.

J.E. Gentle, S.C. Narula, V.A. Sposito (1987) Algorithms for unconstrained L_1 linear regression. In Y. Dodge (ed.)

Statistical data analysis based on the L_1 norm and related methods. North-Holland. 83-94.

J.E. Gentle, V.A. Sposito (1976) On the invariance of certain estimators. Bull. Austral. Math. Soc., vol.14, 405-408.

J.E. Gentle, V.A. Sposito, W.J. Kennedy (1977) On some properties of L_1 estimators. Math. Prog., 12, 139-140.

J.E. Gentle, V.A. Sposito, S.C. Narula (1988) Algorithms for unconstrained L_1 simple linear regression. CSDA, 6(4), 335-340.

W.M. Gentleman (1965) Robust estimation of multivariate location by minimizing p-th power deviations. Ph.D. thesis, Princeton University, New Jersey.

J. Gilsinn, K. Hoffman, R.H.F. Jackson, E. Leyendecker, P. Saunder, D. Shier (1977) Methodology and analysis for comparing discrete L_1 approximation codes., Commun. Stat., B6, 399-413.

F.R. Glahe, J.G. Hunt (1970) The small sample properties of simultaneous equation least absolute estimators vis-a-vis least squares estimators. Econometrica, 38, 742-753.

K. Glashoff, R. Schultz (1979) Uber die genaue Berechnung von besten L_1-approximierenden. J. Approx. Theory 25, 280-293.

S.M. Goldfeld, R.E. Quandt (1981) Econometric modelling with non-normal disturbances. J. of Econometrics, Nov., 17(2), 141-55.

A.A. Goldstein, W. Cheney (1958) A finite algorithm for the solution of consistent linear equations and inequalities and for Tchebycheff approximation of inconsistent linear equations. Pacific J. Math., 8, 415-427.

R. Gonin (1983) A contribution to the solving of nonlinear estimation problems. Ph.D. thesis, University of CapeTown.

R. Gonin (1986) Numerical algorithms for solving nonlinear L_p-norm estimation problems: part I; a first-order gradient algorithm for well-conditioned small residual problems. Comm. Stat. B, 15(3), 801-813.

R. Gonin, A.H. Money (1985a) Nonlinear L_p-norm estimation: part I, on the choice of the exponent, p, where the errors are additive Comm. Stat. A14, 827-840.

R. Gonin, A.H. Money (1985b) Nonlinear L_p-norm estimation: part II, asymptotic distribution of the exponent, p, as a function of the sample kurtosis. Stat. A14, 841-849.

R. Gonin, A.H. Money (1987a) Outliers in physical processes: L_1- or adaptive L_p norm estimation. In Y. Dodge (ed.) Statistical data analysis based on the L_1 norm and related methods. North-Holland. 477-454.

R. Gonin, A.H. Money (1987b) A review of computational methods for solving the nonlinear L_1 norm estimation problem In Y. Dodge (ed.) Statistical data analysis based on the L_1 norm and related methods. North-Holland. 117-130.

R. Gonin, A.H. Money (1987c) Nonlinear L_p-norm parameter estimation. Draft manuscript, Marcel Dekker, New York.

R. Gonin, S.H.C. du Toit (1987) Numerical algorithms for solving nonlinear L_p-norm estimation problem, partK: II- a mixture method for large residual and ill-conditioned problems. Comm. Stat. A16, no. 4.

S. Gross, W.L. Steiger (1979) Least absolute deviation estimates in autoregression with infinite variance. J. Appl. Prob., 16, 104-116.

G. Groucher (1971) Best L_1 and L_∞ approximations. M.Sc. thesis, Birkbeck College, London University, London, England.

M.R. Gupta (1984) Functional forms for estimating the Lorenz curve. Econometrica, 52, 1313-1314.

S.A. Gustafson, K.O. Kortanek, W. Rom (1970) Non-Chebyshevian moment problems. SIAM J. Numer. Anal., vol. 7, no. 3, 335-342.

L. Gyorfi (1987) Density estimation from dependent sample. In Y. Dodge (ed.) Statistical data analysis based on the L_1 norm and related methods. North-Holland. 393-404.

L. Gyorfi, E.C. Van der Meulen (1987) Density-free convergence properties of various estimators of entropy. CSDA, 5(4), 425-436.

E.E. Hagen (1975) The economics of development. Irwin inc, Illinois.

J. Hald (1981a) A 2-stage algorithm for nonlinear L_1 optimization. Rep. no. 81-03, Numerisk Instut. Danmark Tekniske Hojskole, 2800 Lyngby, Denmark.

J. Hald (1981b) A two stage algorithm for linearly constrained nonlinear L_1 optimization. Methods of Oper. Res., 43, 87-103.

J. Hald, K. Madsen (1985) Combined L_p and quasi-Newton methods for nonlinear L_1 optimization. SIAM J. Numer. Anal., 22, no.1, 68-80.

M.L. Hand, V.A. Sposito (1980) Using the least squares estimator in the Chebyshev estimation. Commun. Stat., B9(1), 43-49.

W. Hardle (1987) XploRe, a computing environment for exploratory regression. In Y. Dodge (ed.) Statistical data analysis based on the L_1 norm and related methods. North-Holland. 163-174.

T.E. Harris (1950) Regression using minimum absolute deviations. Am. Statist., 4, 14-15.

H.L. Harter (1974a) The method of least squares and some alternative, I. Int. Stat. Rev., 42, 147-174.

H.L. Harter (1974b) The method of least squares and some alternative, II. Int. Stat. Rev., 42, 235-264.

H.L. Harter (1975a) The method of least squares and some alternative, III. Int. Stat. Rev., 43, 1-44.

H.L. Harter (1975b) The method of least squares and some alternative, IV Int. Stat. Rev., 43, 125-190, 273-278.

H.L. Harter (1975c) The method of least squares and some alternative, V. Int. Stat. Rev., 43, 269-272.

H.L. Harter (1976) The method of least squares and some alternative, VI. Int. Stat. Rev., 44, 113-159.

H.L. Harter (1977) Nonuniqueness of absolute value regression. Comm. Stat. B6, 829-838.

H.L. Harter (1981), Method of least p-th powers. In Encyclopedia of statistical science, 5, 464-467.

A.C. Harvey (1977) A comparison of preliminary estimators for robust regression. JASA, 72, 910-13.

A.C. Harvey (1978) On the unbiasedness of robust regression estimators. Comm. Stat., A7, 779-783.

P. Hattenschwiler (1988) Goal programming becomes most useful using L_1-smoothing functions CSDA, 6(4), 369-384.

W.M. Haussler (1984) Computational experience with an eigen vector algorithm for robust L_p-discrimination. Com. Stat. Q. 1, 233-244.

W.J. Heiser (1987) Correspondence analysis with least absolute residuals CSDA, 5, 337-356.

W.J. Heiser (1988) Multidimensional scaling with least absolute residuals. To appear in H.H. Boc (ed.), Classification and related methods of data analysis, (IFCS'87). North-Holland.

S. Henriksson (1972) On a generalization of L_p-approximation and estimation. Thesis, Dept. of Computer Sci., Lund university, Sweden.

R.W. Hill, P.W. Holland (1977) Two robust alternatives to least squares regression. JASA, 72, 828-833.

C.A.R. Hoare (1961) Algorithm 63 partition; 64, quicksort; and 65, find., Comm. ACM, 4, July, 321-322.

C.A.R. Hoare (1962) Quicksort. Comput. J., 5, 10-15.

C.R. Hobby, J.R. Rice (1965) A moment problem in L_1 approximation. Proc. Amer. Math. Soc., 16, 665-670.

K.L. Hoffman, D.R. Shier (1980a) A test problem generator for discrete linear L_1 approximation problems. ACM Trans. Math. Soft., 6, 587-593.

K.L. Hoffman, D.R. Shier (1980b) A test problem generator for discrete linear L_1 approximation problems. ACM Trans. Math. Soft., 6, 615-617.

W.W. Hogan (1976) Norm minimizing estimation and unbiasedness. Econometrica, vol. 44, no.3, May.

P.W. Holland, R.E. Welsch (1977) Robust regression using iteratively reweighted least-squares. Comm. Stat., A6, 813-827.

L. Horvath (1987) Asymptotic normality of L_p-norms of density estimators. Tech. Rep. series of Lab Res. Stat. Prob., no.3, Carleton University, Ottawa, Canada.

Th.V. Hromadka II, Ch.Ch. Yen, G.F. Pinder (1987) The best approximation method. An introduction. Springer-Verlag, Berlin.

P.J. Huber (1987) The place of the L_1-norm in robust estimation. In Y. Dodge (ed.) Statistical data analysis based on the L_1 norm and related methods. North-Holland. Reprinted in CSDA, 5, 255-262.

C.R. Hunt (1970) Best L_p approximation by certain nonlinear functions. M.Sc. thesis, University of Victoria, University of Victoria, B.C. Canada.

J.G. Hunt, J.M. Dowling, F.R. Glahe (1974) L_1 estimation in small samples with Laplace error distributions Decision Sci., 5, 22-29.

H. Imai, K. Kato, P. Yamamoto (1987) A linear-time algorithm for linear L_1 approximation of points. Tech. Rep. CSCE-87-C30. Dept. of Comp. Sci. and Commun. Engin., Kyushu University 36, Fukuoka 812, Japan.

K. Jajuga (1987) A clustering method based on the L_1-norm. CSDA, 5, 357-371.

K. Jittorntrum, M.R. Osborne (1980) Strong uniqueness and second order convergence in nonlinear discrete approximation. Numer. Math., 34, 439-455.

B. Joe, R. Bartels (1983) An exact penalty method for constrained, discrete linear L_∞ data fitting. SIAM J. Sci. Stat. Comput., vol.4, no.1, 69-84.

L.A. Josvanger, V.A. Sposito (1983) L_1-norm estimates for the simple regression problem. Comm. Stat. B12, 215-21.

J. Jureckova (1983) Trimmed polynomial regression. Commentationes Mathematicae Universitatis Carolinae, 24, 4, 597-607.

J. Jureckova (1984) Regression quantiles and trimmed least squares estimator under a general design. Kybernetika, vol.20, no.5, 345-357.

J. Jureckova, P.K. Sen (1984) On adaptive scale-equivalent M-estimators in linear models. Stat. and Decision supplement issue, no.1, 31-46.

K.R. Kadiyala (1972) Regression with non-Gaussian stable disturbances: some sampling results. Econometrica, July.

N. Kaergard (1987) Estimation criterion, residuals and prediction evaluation. CSDA, 5, 443-450.

S.W. Kahng (1972) Best L_p approximation. Math. of Comp., 26, 505-508.

N.C. Kakwani (1980) Functional forms for estimating the Lorenz curve: a reply. Econometrica, 48, 1063-64.

N.C. Kakwani, N. Podder (1976) Efficient estimation of the Lorenz curve and associated inequality measures from grouped observations. Econometrica 44, 137-148.

L.V. Kantorovich, G.P. Akilov (1964) Functional analysis in normed spaces. Pergamon Press, Oxford England.

L.A. Karlovitz (1970a) An algorithm for the best L_p approximation. J. Approx. Theory. 3, 123-127.

L.A. Karlovitz (1970b) Construction of nearest points in the L_p, even and L_∞ norms. J. Approx. Theory, 3, 123-127.

O.J. Karst (1958) Linear curve fitting using least deviations. JASA, 53, 118-132.

L. Kaufman, P.J. Rousseeuw (1987) Clustering by means of medoids. In Y. Dodge (ed.) Statistical data analysis based on the L_1 norm and related methods. North-Holland. 405-416.

Y. Kawara (1979) Straight line fitting by minimizing the sum of absolute deviations. J. of the Japan Stat. Soc., 9, 47-64.

J.E. Kelley (1958) An application of linear programming to curve fitting. SIAM J. Appl. Math., 6, 15-22.

J.H.B. Kemperman (1984) Least absolute value and median polish. In Y.L. Tong (ed.), Inequalities in statistics and probability (IMS Lecture notes monograph series, vol.5), Inst. of Math. Stat., Hayward, CA, 84-113.

J.H.B. Kemperman (1987) The median of a finite measure on a Banach space. In Y. Dodge (ed.) Statistical data analysis based on the L_1 norm and related methods. North-Holland. 217-230.

M. Kendall, A. Stuart (1977) The advanced theory of statistics. vol.1, Charles Griffin & Co., London.

W.J. Kennedy, J.E. Gentle (1977) Examining rounding error in least absolute values regression computations. Comm. Stat., B6, 415-420.

W.J. Kennedy, J.E. Gentle (1978) Comparisons of algorithms for minimum L_p norm linear regression. Proc. of Computer Sci. and Stat., Tenth Annual Symposium on Interface. D. Hogben (ed.), U.S. government printing office. Washington D.C., 373-378.

W.J. Kennedy, J.E. Gentle (1980) Statistical computing. New York, Marcel Dekker.

W.J Kennedy, J.E. Gentle, V.A. Sposito (1977) Comparisons of algorithms for L_1 estimation in the linear model. Paper presented at Midwestern Regional Meeting of IMS, Madison, WI.

(Available from the second author).

W.J Kennedy, J.E. Gentle, V.A. Sposito (1977) A computer oriented method for generating test problems for L_1 regression. Comm. Stat., B6, 21-27.

B. Kim ( ) L_p norm estimation procedures and an L_1 norm algorithm for unconstrained and constrained estimation for linear models. VPI & SU, Blacksurg, VA 24061 USA.

E.A. Kiountouzis (1971) Optimal L_p approximation techniques and data analysis. Bull. Soc. Math. Greece, 12, 191-206.

E.A. Kiountouzis (1972) Mathematical programming and best linear L_p approximation. Extrait du Bull. de la Soc. Math. de Grece Novelle Serie, Tom 13, Fassc. 1, 46-57.

E.A. Kiountouzis (1973) Linear programming techniques in regression analysis. Appl. Stat., 22, 69-73.

D. Klingman, J. Mote (1982) Generalize network approaches for solving least absolute value and Tchebycheff regression problems. TIMS studies in the Mana. Sci., 19.

J. Kmenta (1986) Elements of econometrics. MacMillan, New York.

R. Koenker (1987) A comparison of asymptotic testing methods for L_1-regression. In Y. Dodge (ed.) Statistical data analysis based on the L_1 norm and related methods. North-Holland. 287-295.

R. Koenker, G. Bassett, Jr. (1978) Regression quantile. Econometrica, vol. 46, no. 1 (Jan.).

R. Koenker, G. Bassett, Jr. (1982a) Test of linear hypothesis and L_1 estimation. Econometrica, vol.50, no. 6, 1577-1583.

R. Koenker, G. Bassett, Jr. (1982b) Robust tests for heteroskedasticity based on the regression quantiles. Econometrica, vol. 50, no. 1, Jan., 43-61.

R. Koenker, G. Bassett, Jr. (1984) Four (pathological) examples in asymptotic statistics. The Amer. Statistician, Aug., vol.38, no.3, 209-212.

R. Koenker, G. Bassett, Jr. (1985) On Boscovich's estimator. Annals of Stat., 13, 1625-1628.

W.W. Kotiuga (1982) Power system state estimation using least absolute value techniques. Ph.D. thesis, University of Waterloo.

W.W. Kotiuga, M. Vidyasagar (1982) Bad data rejection properties of weighted least absolute value techniques applied to static state estimation. IEEE Trans. on Power Apparatus and Systems. PAS-101, 844-853.

B.R. Kripke, T.J. Rivlin (1965) Approximation in the metric of $L_1(X,u)$. Trans. Amer. Math. Soc., 119, 101-22.

K. Kveton (1987) Method of averages as an alternative to L_1- and L_2-norm methods in special linear regression problems. CSDA, 5, 407-414.

P.S. Laplace (1793) Sur quelques points du system du monde. Memoires de l'Academie Royale des Science de Paris. Annee 1789, 1-87. Reprinted in Oeuvres completes de Laplace II. Paris, Gauthier-Villars, 1985, 477-558.

P.S. Laplace (1799) Traite des mecanique celeste, 2. Paris; J.B.M. Depart. Reprinted as oeuvres completes de Laplace, 2. Paris; Gauthier-Villars 1878, 116-165.

P.S. Laplace (1812) Theorie analytique des probabilites, Mme courcier Paris 1820 Reprinted in his oeuvres, vol.7, Imprimerie Royale, Paris, 1847, and Gauthier-Villars et fils, Paris 1886.

P.S. Laplace (1818) Duexieme supplement to Laplace (1812).

J.L. Larson, A.H. Sameh (1980) Algorithms for round of error analysis a relative error approach. Computing 24, 275-297.

K.D. Lawrence, D.R. Shier (1981) A comparison of least square and least absolute deviation regression models for estimation Weibull parameters. Comm. Stat., B10, 315-326.

C.L. Lawson (1961) Contribution to the theory of linear least maximum approximations. Ph.D. thesis, University of California, Los Angeles, California.

E. Lazarski (1975a) Approximation of continuous functions in the space L_1. Automatika, 487, 85-93.

E. Lazarski (1975b) The approximation of the continuous function by the polynomials of power functions in L_1 space. Automatika, 487, 95-106.

E. Lazarski (1975c) On the necessary conditions of the uniqueness of approximation by the polynomials of power functions in L_1 space. Automatika, 487, 107-117.

E. Lazarski (1977) Approximation of continuous functions by exponential polynomials in the L_1 space. Automatika, 598, 82-87.

M.G. Lejeune, P. Sarda (1988) Quantile regression: a non parametric approach. CSDA, 6(3) 229-240.

J.T. Lewis (1969) Approximation with convex constraints. Doctoral thesis, Brown university, Providence, R.I.

J.T. Lewis (1970) Computation of best one-sided L_1 approximation. Math. Comp., 24, 529-536.

R.F. Love (1974) The dual of a hyperbolic approximation to the generalized constrained multi-facility location problem with L_p distances Manag. Sci., vol. 21, 22-23.

R.F. Love, J.G. Morris (1975) The use of nonlinear programming for solving facility location problems involving L_p distances using convex programming. Oper. Res., vol. 23, no.3, 581-588.

G.S. Maddala (1977) Econometrics. McGraw-Hill.

K. Madsen (1975) An algorithm for minimax solution of overdetermined systems of nonlinear equations. J. Inst. Math. and Appl., 321-328.

K. Madsen (1985) Minimization of non-linear approximation functions. Copenhagen.

B. Mandelbrot (1960) The Pareto-Levy law and the distribution of income. Inter. Econ. Rev., 1, 79-106.

B. Mandelbrot (1961) Stable Paretian random functions and multiplicative variation of income. Econometrica, 29, 517-543.

B. Mandelbrot (1963) New methods in statistical economics. J. of Political Economy, Oct.,421-440.

A. Marazzi (1987) Solving bounded influence regression with ROBSYS. In Y. Dodge (ed.) Statistical data analysis based on the L_1 norm and related methods. North-Holland. 145-163.

A. Marazzi (1988) Algorithms for the computation of weights in bounded influence regression. CSDA 6(3), 251-276.

A. Marazzi, A. Randriamiharisoa (1985) ROBETH-ROBSYS: a software for robust statistical computing. Document no. 0,1,2,3,4,6. Institut Universitaire de Medecin Sociale et Preventive Lausanne, Switzerland.

J.S. Marron (1987) What does optimal bandwidth selection mean for non parametric regression estimation. In Y. Dodge (ed.)

Statistical data analysis based on the L_1 norm and related methods. North-Holland. 379 392.

C.L. Mathieu (1816) Sur les experiences du pendule, faites par les navigateurs espagnol, en differens points du globe. Connaissance des tems, 314-332.

J.W. McKean, R.M. Schrader (1984) A comparison of the methods for studentizing the sample median. Comm. Stat., B 13(16), 751-773.

J.W. McKean, R.M. Schrader (1987) Least absolute errors analysis of variance. In Y. Dodge (ed.) Statistical data analysis based on the L_1 norm and related methods. North-Holland. 297-306.

J.W. McKean, G.L. Sievers (1987) Coefficients of determination for least absolute deviation analysis. Stat. and Prob. Letters, 5, 49-54.

R.A. McLean, G.A. Watson (1980) Numerical methods for nonlinear discrete L_1 approximation problems. In L. Collatz, G. Meinardus, H. Werner (eds.) Numerical methods of approximation theory. ISNM 52, Birkhauser Verlag, Basel.

C.R. McConnell (1987) On computing a best discrete L_1 approximation using the method of vanishing Jacobians. CSDA, 5, 277-288.

G.F. McCormick, V.A. Sposito (1975) A note on L_1 estimation based on the median positive quotient. Appl. Stat., 24, 347-350.

G.F. McCormick, V.A. Sposito (1976) Using the L_2-estimator in L_1-estimation. SIAM J. Numer. Anal., 13, 337-343.

N. Megiddo, A. Tamir (1983) Finding least-distances line. SIAM J. Alg. Disc. Meth., 4, no. 2, 207-211.

J. Meier (1987) A fast algorithm for clusterwise linear absolute deviations regression. OR Spektrum, 9, 187-189.

M.S. Meketon (1986) Least absolute value regression. Work. Pap., AT&T Bell Laboratories, Holmdel, N.J.

A. Melaku, G. Sadasivan (1987) L_1-norm and other methods for sample allocation in multivariate stratified surveys. CSDA, 5, 415-424.

J.A. Menendez, B. Salvador (1987) An algorithm for isotonic median regression. CSDA, 5, 399-406.

G. Merle, H. Spath (1974) Computational experiences with discrete L_p-approximation. Computing, 12, 315-321.

J.R. Meyer, R.R. Glauber (1964) Investment decisions, Economic forecasting and public policy. Harvard Business School Press, Cambridge, Massachusetts.

J. Militky, J. Cap (1987) Application of Bayes approach to adaptive L_p nonlinear regression. CSDA, 5, 381-390.

J.S.B. Mitchell (1987) Shortest rectilinear paths among obstacles. School of Oper. Res. and Ind. Engin. College of Engin. Cornell University. Ithaca, New York, 14853.

M.J. Mojarrad (1977) The application of comparative Monte Carlo methods to econometrics: an efficient Monte Carlo study of finite sample properties of iterative instrumental variables estimation. Ph.D. Diss., University of Pennsylvania.

B. Mond, M. Schechter (1976) A programming problem with an L_p norm in the objective function. J. Austral. Math. Soc., Ser., B, 19, 333-342.

A.H. Money, J.F. Affleck-Graves, M.L. Hart (1978a) A review of some alternatives to least squares regression. Tech Rep.

no. ALS-1, Sep., University of CapeTown, South Africa.

A.H. Money, J.F. Affleck-Graves, M.L. Hart (1978b) Least squares and some alternative: a simulation study. Tech. Rep. ALS-2. University of CapeTown, South Africa.

A.H. Money, J.F. Affleck-Graves, M.L. Hart, G.D.I. Barr (1982) The linear regression model: L_p norm estimation and the choice of p. Comm. Stat., 11, 89-109.

R.M. Moroney (1961) The Haar problem in L_1. Proc. Amer. Math. Soc., 12, 793-795.

J.G. Morris, W.a Verdini (1979) Minisum L_p distance location problems solved via a perturbed problem and Weiszfeld's algorithm. Oper. Res., 27, 1180-1188.

J. Munoz Perez, A. Fernandez Palacin (1987) Estimating the quantile function by Bernstein polynomials. CSDA, 5, 391-398.

V.I. Mudrov, V.L. Kushko, V.I. Mikhailov, E. Osvitskii (1968) Some experiments on the use of the least-modul: method in processing orbital data. Cosmic Res., 6, 421-431.

W. Murray, M. Overton (1981) A projected Lagrangian algorithm for nonlinear L_1 optimization. SIAM J. Sci. Stat. Comput., 207-224.

S.C. Narula (1987) The minimum sum of absolute errors regression. J. Quality Tech. 19, 37-45.

S.C. Narula, J.F. Wellington (1977a) An algorithm for the minimum sum of weighted absolute errors regression. Comm. Stat., B(6), 341-352.

S.C. Narula, J.F. Wellington (1977b) Prediction, linear regression and minimum sum of relative errors. Technometrics, 19, 185-190.

S.C. Narula, J.F. Wellington (1977c) AS108, multiple linear regression with minimum sum of absolute error. Appl. Stat., 26, 106-111.

S.C. Narula, J.F. Wellington (1979) Selection of variables in linear regression using the minimum sum of weighted absolute errors criterion. Technometrics, 21, no.3 Aug.

S.C. Narula, J.F. Wellington (1982) The minimum sum of absolute errors regression, a state of the art survey. Inter. Stat. Rev., 50, 317-326.

S.C. Narula, J.F. Wellington (1983) Selection of variables in linear regression, a pragmatic approach. J. of Stat. Comput. and Simul., 17, 159-172.

S.C. Narula, J.F. Wellington (1985) Interior analysis for the minimum sum of absolute errors regression. Technometrics, 27, 181-188.

S.C. Narula, J.F. Wellington (1987) An efficient algorithm for the MSAE and MMAE regression problems. Work. Pap., Virginia CommonWealth University, Richmond, VA 23284.

H. Nikaido (1970) Introduction to sets and mappings in modern economics. North Holland, Amsterdam.

H. Nyquist (1980) Recent studies on L_p-norm estimation. Ph.D. thesis, University of Umea, Sweden.

H. Nyquist (1983) The optimal L_p-norm estimator in linear regression models. Comm. Stat. A12, 2511-2524.

H. Nyquist (1988) Least orthogonal absolute deviations. CSDA, 6(4), 361-368.

H. Nyquist, A. Westlund (1977) L_1-versus L_2-norm estimation in interdependent systems when residual distributions are stable. Dept. of Stat., University of Umea Presented at

European Meeting of the Econometric Soc., Vienna, 5-9 Sep.

W. Oberhofer (1982) The consistency of nonlinear regression minimizing the L_1 norm. Ann. of Stat., 10, 316-319.

W. Oettli (1975) Symmetric duality, and a convergent subgradient method for discrete linear, constrained approximation problems with arbitrary norms appearing in the objective function and in the constraints. J. Approx. Theory, 14, 43-50.

M.R. Osborne (1980) An algorithmic approach to nonlinear approximation problems. Approx. Theory III, 705-710.

M.R. Osborne (1985) Finite algorithms in optimization and data analysis. Wiley, Chichester.

M.R. Osborne (1987) The reduced gradient algorithm. In Y. Dodge (ed.) Statistical data analysis based on the L_1 norm and related methods. North-Holland. 95-108.

M.R. Osborne, S.A. Pruess, R.S. Womersley (1986) Concise representation of generalized gradients. J. of Austra. Math. Soc., Ser. B, 28, 57-74.

M.R. Osborne, G.A. Watson (1967) On the best linear Chebyshev approximation. Computer J., 10, 172-177.

M.R. Osborne, G.A. Watson (1971) On an algorithm for discrete nonlinear L_1 approximation. Computer J., 14, 184-188.

M.R. Osborne, G.A. Watson (1978) Nonlinear approximation problems in vector norms. In Dundee, G.A. Watson (eds.) Numerical analysis. Springer Verlag.

M.R. Osborne, G.A. Watson (1985) An analysis of the total approximation problem in separable norms, and an algorithm for the total L_1 problem. SIAM J. Sci. Stat. Comp., 6, 410-424.

M. Overton (1982) Algorithms for nonlinear L_1 and L_∞ fitting. In M.J.D. Powell (ed.) Nonlinear optimization. Academic Press, London, 91-101

R.M. Oveson (1968) Regression parameter estimation by minimizing the sum of absolute errors. Doctoral dissertation Harvard university, Cambridge, Massachusetts.

H.J. Paarsch (1984) A Monte Carlo comparison of estimates for censored regression models. J. of Econometrics, 24, 197-213.

M.J. Panik (1976) Classical optimization: foundation and extensions. North-Holland, Amsterdam.

U. Peters, C. Willms (1983) Up- and down-dating procedures for linear L_1 regression. OR Spektrum 5, 229-239.

R.C. Pfaffenberger, J.J. Dinkel (1978) Absolute deviations curve fitting: an alternative to least squares. In H.A. David (ed.) Contributions to survey sampling and applied statistics. Academic Press, New York, 279-294.

P. Pilibossian (1987) A direct solving algorithm for a linear regression according to L_1-norm criteria. Work. Pap., L.S.T.A. Universite, Paris VI

M.A. Porter, D.J. Winstanley (1979) Remark ASR29. Remarks on AS110: L_p norm fit of a straight line. Appl. Stat., 28, 112-113.

S. Portnoy (1987) Using regression fractiles to identify outliers. In Y. Dodge (ed.) Statistical data analysis based on the L_1 norm and related methods. North-Holland, 345-356.

J.L. Powell (1983) The asymptotic normality of two-stage least absolute deviation estimators. Econometrica, 51, 1569-1575.

J.L. Powell (1984) Least absolute deviations estimation for the censored regression model. J. of Econometrics, 25, 303-325.

J.L. Powell (1986) Censored regression quantiles. J. of Econometrics, 32, 143-155.

M.J.D. Powell, Y. Yuan (1984) Conditions for super linear convergence in L_1 and L_∞ solutions of overdetermined nonlinear equations. IMAJ Num. Anal., 4, 241-251.

B. Prochazka (1988) Regression quantiles and trimmed least squares estimator in nonlinear regression model. CSDA, 6(4), 385-392.

R. Prony (1804) Recherches physico-mathematiques sur la theorie des eaus courantes. Paris, l'imprimerie imperiale.

V. Ptak (1958) On approximation of continuous functions in the metric $\int_a^b |x(t)|dt$ Czechoslovak Math. J. 8(83), 267-273.

P. Rabinowitz (1968) Application of linear programming to numerical analysis. SIAM Rev., 10, 121-159.

P. Rabinowitz (1970) Mathematical programming and approximation. In A. Talbot (ed.) Approximation Theory. Academic Press, 217-231.

A. Ralston, P. Rabinowitz (1985) A first course in numerical analysis. Wiley, New York.

J.O. Ramsay (1977) A comparative study of several robust estimates of slopes, intercept, and scale in linear regression. JASA, 72, 608-615

M.R. Rao, V. Srinivasan (1972) A note on Sharpe's algorithm for minimum sum of absolute deviations in a simple regression problem. Manag. Sci., 19, 222-225.

R.H. Rasche, J. Gaffney, A.Y.C. Koo, N. Obst (1980) Functional forms for estimating the Lorenz curve. Econometrica, 48, 1061-1062.

H.D. Ratliff, J.C. Picard (1978) A cut approach to rectilinear distance facility location problem. Oper. Res., 26, 422-433.

W. Rey (1975) On the least p_{th} power methods in multiple regressions and location estimations. BIT, 15, 174-185.

E.C. Rhodes (1930) Reducing observations by the method of minimum deviations. Philo. Mag., 7^{th} series, 9, 974-92.

J.R. Rice (1964a) On computation of L_1 approximations by exponentials, rationals, and other functions. Math. Comp., 18, 390-396.

J.R. Rice (1964b) On nonlinear L_1 approximation. Arch. Rational Mech. Anal., 17 61-66.

J.R. Rice (1964c) The approximation of functions, vol. I, linear theory. Reading Mass:, Addison-Wesley.

J.R. Rice (1969) The approximation of functions, vol. II, linear theory. Reading Mass:, Addison-Wesley.

J.R. Rice (1985) Numerical methods, software, and analysis. McGraw-Hill, ch. 11.

J.R. Rice, J.S. White (1964) Norms for smoothing and estimation. SIAM Rev., 6, 243-256.

P.D. Robers, A. Ben-Israel (1969) An interval programming algorithm for discrete linear L_1 approximation problem. J. Approx. Theory, 2, 323-336.

P.D. Robers, S.S. Robers (1973) Algorithm 458: discrete linear L_1 approximation by interval linear programming. Comm. ACM, 16, 629-633.

E. Ronchetti (1987) Bounded influence in regression: a review. In Y. Dodge (ed.) Statistical data analysis based on the L_1 norm and related methods. North-Holland, 65-80.

A.E. Ronner (1977) P-norm estimators in a linear regression model. Doctoral thesis, Rijkuniversiteit te Groningen.

A.E. Ronner (1984) Asymptotic normality of p-norm estimators in multiple regression. Z. Wahrschein lickeitstheorie verw. Gebiete 66, 613-620.

G. Roodman (1974) A procedure for optimal stepwise MSAE regression analysis. Oper. Res., 22, 393-399.

B. Rosenberg, D. Carlson (1977) A simple approximation of the sampling distribution of least absolute residuals regression estimates. Comm. Stat., B6, 421-437.

R. Rossi, H.D. Brunk (1987) L_1 and L_2 cross-validation for density estimation with special reference to orthogonal expansions. Tech. Rep. 120, Dept. of Stat., Oregon State University.

R. Rossi, H.D. Brunk (1988) L_1 and L_2 cross-validation for density estimation with special reference to orthogonal expansions. CSDA, 6(3), 203-228.

P.J. Rousseeuw (1984) Least median of squares regression. J. Amer. Stat. Asso., 79, 871-80.

P.J. Rousseeuw (1987) An application of L_1 to astronomy. In Y. Dodge (ed.) Statistical data analysis based on the L_1 norm and related methods. North-Holland, 437-446.

P.J. Rousseeuw, A. Leroy (1987) Robust regression and outlier detection. Wiley-Interscience, New York.

A.N. Sadovski (1974) AS74: L_1-norm fit of a straight line. Appl. Stat. 23, 244-248.

A.K.Md.E. Saleh, P.K. Sen (1987) On the asymptotic distributional risk properties of pre-test and shrinkage L_1 estimators. CSDA, 5, 289-300.

J.P. Schellhorn (1987) Fitting data through homotopy methods In Y. Dodge (ed.) Statistical data analysis based on the L_1 norm and related methods. North-Holland. 131-138.

E.J. Schlossmacher (1973) An iterative technique for absolute deviations curve fitting. JASA 68, 857-865.

R.M. Schrader, J.W. McKean (1987) Small sample properties of least absolute errors analysis of variance. In Y. Dodge (ed.) Statistical data analysis based on the L_1 norm and related methods. North-Holland. 307-322.

R.S. Scowen (1965) Algorithm 271 Quickersort. Commun. ACM, 8, 669-670.

E. Seneta (1983) The weighted median and multiple regression. Austral. J. Stat., 25(2), 370-377.

E. Seneta, W.L. Steiger (1984) A new LAD curve-fitting algorithm: slightly overdetermined equation system in L_1. Discrete Applied Math., 7, 79-91.

D. Shanno, R.L. Weil (1970) Linear programming with absolute value functionals. Oper. Res., 19, 120-124.

W.F. Sharpe (1971) Mean-absolute deviation characteristic lines for securities and portfolios. Manag. Sci., 18, B1-B13.

S.J. Sheather (1986) A finite sample estimate of the variance of the sample median. Stat. and Prob. Letter., 4, 337-342.

S.J. Sheather (1987) Assessing the accuracy of the sample median: estimated standard errors versus interpolated confidence intervals. In Y. Dodge (ed.) Statistical data analysis based on the L_1 norm and related methods. North-Holland. 203-216.

S.J. Sheather, J.W. McKean (1987) A comparison of testing and confidence interval methods for the median. Stat. and Prob. Letter, 6, 31-36.

H.D. Sherali, B.O. Skarpness, B. Kim (1987) An assumption-free convergence analysis for a perturbation of the scaling algorithm for linear programs, with application to the L_1 estimation problem. Dept. of Ind. Engin. and OR, Virginia Polytechnic Inst. and State University, Blacksburg, Virginia.

O.B. Sheynin (1973) R.J. Boscovich's work on probability. Archive for history of exact sciences, vol. 9, 306-324, and vol. 28, 173.

D.R. Shier, C.J. Witzgall (1978) Norm approximation problems and norm statistics., J. Res. Nat. Bur. Standards, 83, 71-74.

O. Shisha (1974) A remark on an algorithm for best L_p approximation. J. Approx. Theory, 11, 283-284.

R.I. Shrager, E. Hill (1980) Nonlinear curve-fitting in the L_1 and L_∞ norms. Math. Comput., 34, 529-541.

A.F. Siegel (1983) Low median and least absolute residual analysis of two-way tables. JASA, 78, 371-374.

R.L. Sielken, H.O. Hartley (1973) Two linear programming algorithms for unbiased estimation of linear models. JASA, 68, 639-.

H.A. Simon (1955) On a class of skew distribution functions. Biometrika, 42, 425-440. Reprinted in H.A. Simon (1957) Models of man. New York, Wiley.

R.R. Singleton (1940) A method for minimizing the sum of absolute values of deviations. Annals of math. Stat., 11, 301-310.

R.S. Singleton (1969) Algorithm 347 Sort. Comm. ACM, 12, 185-186.

M.G. Sklar, R.D. Armstrong (1982) Least absolute value and Chebyshev estimation utilizing least squares results. Math. Prog., 24, 346-352.

V.K. Smith, T.W. Hall (1972) A comparison of maximum likelihood versus BLUE estimators. Rev. Econ. Stat., 54, 186-190.

S.A. Soliman, G.S. Christensen, A. Rouhi (1988) A new technique for curve fitting based on minimum absolute deviations. CSDA, 6(4), 341-352.

D.L. Souvaine, J.M. Steele (1987) Time- and space-efficient algorithms for least median of squares regression. JASA, 82, no. 399, 794-801.

H. Spath (1976) L_1 cluster analysis. Computing, 16, 379-387.

H. Spath (1982) On discrete linear orthogonal L_p-approximation. Numerische Analysis, ZAMM 62, T354-T355.

H. Spath (1985) Cluster dissection and analysis. Horwod, Chichester.

H. Spath (1986a) Clusterwise linear least squares versus least absolute deviations regression, a numerical comparison for a case study. In W. Gaul, M. Schader (eds.) Classification as a tool of research. Elsevier, Amsterdam.

H. Spath (1986b) Orthogonal least squares fitting with linear manifolds. Numer. Math., 48, 441-445.

H. Spath (1986c) Algorithm, Clusterwise linear least absolute deviations regression. Computing, 37, 371-378.

H. Spath (1987) Using the L_1 norm within cluster analysis. In Y. Dodge (ed.) Statistical data analysis based on the L_1 norm and related methods. North-Holland. 427-434.

H. Spath, G.A. Watson (1987) On orthogonal linear L_1 approximation. Numer. Math., 51, 531-543.

M.R. Spiegel (1968) Mathematical handbook. Schaum's outline series, McGraw-Hill, New York.

V.A. Sposito (1976) A remark on algorithm AS74, L_1 norm fit of a straight line. Appl. Stat., 25, 96-97.

V.A. Sposito (1982) On the unbiasedness of the L_p-norm estimators. JASA, 77, 652-653.

V.A. Sposito (1987a) On median polish and L_1 estimators. CSDA, 5, 155-162.

V.A. Sposito (1987b) Some properties of L_p-estimators in robust procedures. Marcel Dekker. In print.

V.A. Sposito, M.L. Hand (1980) Optimal L_p estimators for symmetric distributions. Proc. of ASA, Stat. Comp. Sec.

V.A. Sposito, M. Hand, G. McCormick (1977) Using an approximate L_1 estimator. Comm. Stat., B6, 263-268.

V.A. Sposito, M.L. Hand, B. Skarpness (1983) On the efficiency of using the sample kurtosis in selecting optimal L_p estimators. Comm. Stat., B12, 265-272.

V.A. Sposito, W.J. Kennedy, J.E. Gentle (1977) AS110: L_p norm fit of a straight line. Appl. Stat., 26, 114-118.

V.A. Sposito, W.J. Kennedy, J.E. Gentle (1980) Useful generalized properties of L_1 estimators. Comm. Stat., A9, 1309-1315.

V.A. Sposito, G.F. McCormick, W.J. Kennedy (1975) L_1 estimation strategies based on the simplex algorithm. In Proc. of the eighth symposium on the interface, J.W. France (ed.) Health science computing facility. UCLA, Los Angeles.

V.A. Sposito, W.C. Smith (1976) On a sufficient and necessary condition for L_1 estimation. Appl. Stat., 25, 154-157.

V.A. Sposito, W.C. Smith, G.F. McCormick (1978) Minimizing the sum of absolute deviations. J. of Appl. Stat. and Econometrics, Vandenhoeck and Ruprecht in Gottingen and Zurich series 12.

V.A. Sposito, M. Tvejte (1984) The estimation of certain parameters used in L_1 interface. Proc. of Stat. Comp. Sec. of ASA, 267-270.

K. Spyropoulos, E. Kiountouzis, A. Young (1973) Discrete approximation in the L_1 norm. Comp. J., 16, 180-186.

V.P. Sreedharan (1969) Solution of overdetermined linear equation which minimize error in an abstract norm. Numer. Math., 13, 146-151.

V.P. Sreedharan (1971) Least squares algorithms for finding solutions which minimize error in an abstract norm. Numer. Math., 17, 387-401.

G. Stangenhaus (1987) Bootstrap and interface procedures for L_1 regression. In Y. Dodge (ed.) Statistical data analysis based on the L_1 norm and related methods. North-Holland. 323-332.

G. Stangenhaus, S.C. Narula (1987) Inference procedures for the L_1 regression. Work. Pap., Universidade Estadual de Compainas, Brasil.

J.M. Steele, W.L. Steiger (1986) Algorithms and complexity for least median of squares regression. Discrete Appl. Math., 14, 39-100.

E. Stiefel (1960) Note on Jordan elimination, linear programming and Tchebycheff approximation. SIAM J. Numer. Math., 2, 1-17.

J. Steindl (1965) Random processes and growth of the firms. London, Griffin.

W.L. Steiger (1980) Linear programming via L_1 curve fitting beats simplex. Abstracts, AMS, 80T-C26, 385-386.

W. Steiger, P. Bloomfield (1980) Remark on a paper of Narula and Wellington. Technometrics, 22, 450.

S.M. Stigler (1981) Gauss and invention of least squares. Annals of Stat., 9, 465-474.

S.M. Stigler (1984) Studies in the history of probability and statistics XL, Boscovich, Simpson and a 1760 manuscript note on fitting a linear relation. Biometrica, 71, 3, 615-620.

J. Svanberg (1805) Exposition des operations faites en lappnie pour la determination d'un arc du meridien en 1801, 1802 et 1803,... Stockholm.

T. Taguchi (1972a) On the two-dimensional concentration surface and extensions of concentration coefficient and Pareto distribution to the two dimensional case-I. Annals of the Inst. of Stat. Math., vol. 24, no.2, 355-381.

T. Taguchi (1972b) On the two-dimensional concentration surface and extensions of concentration coefficient and Pareto distribution to the two dimensional case-II. Annals of the Inst. of Stat. Math., vol. 24, no.3, 599-619.

T. Taguchi (1972c) Concentration polyhedron, two dimensional concentration coefficient for discrete type distribution and some new correlation coefficients etc. The Inst. of Stat. Math., 77-115.

T. Taguchi (1973) On the two-dimensional concentration surface and extensions of concentration coefficient and Pareto distribution to the two dimensional case-III. Annals of the Inst. of Stat. Math., vol. 25, no.1, 215-237.

T. Taguchi (1974) On Fechner's thesis and statistics with norm p. Ann. of the Inst. of Stat. Math., vol. 26, no.2, 175-193.

T. Taguchi (1978) On a generalization of Gaussian distribution. Ann. of the Inst. of Stat. Math., vol. 30, no.2, A, 211-242.

T. Taguchi (1981) On a multiple Gini's coefficient and some concentrative regressions. Metron, vol. XXXIX - N.1-2, 5-98.

T. Taguchi (1983) Concentration analysis of bivariate Paretoan distribution. Proc. of the Inst. of Stat. Math., vol. 31, no.1, 1-32.

T. Taguchi (1987) On the structure of multivariate concentration. Submitted to the First International Conference on Statistical Data Analysis Based on the L_1 Norm and Related Methods, Neuchatel, Switzerland.

T. Taguchi (1988) On the structure of multivariate concentration - some relationships among the concentration surface and two variate mean difference and regressions. CSDA, 6, 307-334.

A. Takayama (1974) Mathematical economics. The Dryden Press, Illinois.

L.D. Taylor (1974) Estimating by minimizing the sum of absolute errors. In P. Zarembka (ed.) Frontiers in econometrics. Academic Press.

H.H. Taylor, S.C. Banks, J.F. McCoy (1979) Deconvolution with the L_1 norm. Geophysics, 44, 39-52.

H. Theil (1965) The analysis of disturbances in regression analysis. JASA, 67, 1067-1079.

H. Theil (1971) Principles of econometrics. Wiley.

A Tishler, L. Zang (1982) An absolute deviations curve fitting algorithm for nonlinear models. TIMS studies in Manag. Sci., 19.

D.S. Tracy, K.A. Khan (1987) MRPP tests in L_1 norm. CSDA, 5, 373-380.

E. Trauwaert (1987) L_1 in fuzzy clustering. In Y. Dodge (ed.) Statistical data analysis based on the L_1 norm and related methods. North-Holland. 417-426.

J.W. Tukey (1977) Exploratory data analysis. Reading, Mass. Addison-Wesley.

H.H. Turner (1887) On Mr. Edgeworth's method of reducing observations relating to several quantities. Phil. Mag. (5th series), 24, 466-470.

K.H. Usow (1967a) On L_1 approximation: computation for continuous functions and continuous dependence. SIAM J. of Numer. Anal., 4, 70-88.

K.H. Usow (1967b) On L_1 approximation: computation for discrete functions and discretization effect. SIAM J. Numer. Anal., 4, 233-244.

I. Vajda (1987) L_1-distances in statistical inference: comparison of topological, functional and statistical properties. In Y. Dodge (ed.) Statistical data analysis based on the L_1 norm and related methods. North-Holland. 177-184.

C.W. Valentine, C.P. Van Dine (1963) An algorithm for minimax polynomial curve fitting for discrete data. J. ACM, 10, 283-290.

J.F. Van Beeck-Calkoen (1816) Ver de theoric der Gemiddelde Waardij. Verhandlingen der K. Nederlandandsch Instituut Can Wetenschappen, 2, 1-19.

L. Veidinger (1960) On the numerical determination of the best approximation in the Chebychev sense. Numer. Math., 2, 1-17.

B.A. Von Lindenau (1806) Uber den Gebrauch der Gradmessungen zur bestimmung der gestalt der erde. Monatliche correspondenz zur befar derung der Erd-und Himmels-kunde, 14, 113-158.

B.Z. Vulikh (1976) A brief course in the theory of functions of a real variable. Mir Publishers, Moscow.

H.M. Wagner (1959) Linear programming technique for regression analysis. JASA, 54, 202-212.

H.M. Wagner (1962) Nonlinear regression with minimal assumption. JASA, 57, 572-578.

G.A. Watson (1973) On the best linear Chebyshev approximation. J. Approx. Theory, 7, 48-58.

G.A. Watson (1973) The calculation of best linear one-side L_p approximations. Math. Comp. 27, 607-620.

G.A. Watson (1977) On two methods for discrete L_p-approximation. Computing, 18, 263-266.

G.A. Watson (1978) A class of programming problems whose objective function contains a norm. J. approx. Theory, 23, 401-411.

G.A. Watson (1980) Approximation theory and numerical methods. Wiley, New York.

G.A. Watson (1981) An algorithm for linear L_1 approximation of continuous functions. IMA J. Num. Anal., 1, 157-167.

G.A. Watson (1982a) A globally convergent method for (constrained) nonlinear continuous L_1 approximation problems. In Numerical methods of approximation theory. ISNM59, Birkhauser Verlag.

G.A. Watson (1982b) Numerical methods for linear orthogonal L_p approximation. IMA J. of Numer. Anal., 2, 275-287.

G.A. Watson (1984a) Discrete L_1 approximation by rational functions. IMA J. Num. Anal., 4, 275-288.

G.A. Watson (1984b) The numerical solution of total L_p approximation problems. In D.F. Griffiths (ed.) Numerical analysis. Dundee 1983, Lecture notes in mathematics, 1066, Springer Verlag, 221-238.

G.A. Watson (1985a) On the convergence of eigenvector algorithms for robust L_p-discrimination. Comp. Stat. Quart., 4, 307-314.

G.A. Watson (1985b) On a class of algorithms for total approximation. J. Approx. Theory, 45, no.3, 219-231.

G.A. Watson (1986) Methods for best approximation and regression problems. Rep. NA/96, Dept. of Math. Sci., University of Dundee, DD1 4hn, Scotland, UK.

G.A. Watson (1987) Data fitting by sums of exponentials using the L_1 norm. Inter. Series of Numer. Math., 81, 246-261.

J.F. Wellington, S.C. Narula (1981) Variable selection in multiple linear regression using the minimum sum of weighted absolute errors criterion. Comm. Stat., B10, 641-648.

J.F. Wellington, S.C. Narula (1984) An algorithm for regression quantiles. Comm. Stat., B13(5), 683-704.

A.H. Welsh (1987) Kernel estimates of the sparsity function. In Y. Dodge (ed.) Statistical data analysis based on the L_1 norm and related methods. North-Holland. 369-378.

G.O Wesolowsky (1981) A new descent algorithm for least absolute value regression problem. Comm. Stat., B10, 479-491.

G.O. Wesolowsky, R.F. Love (1971) The optimal location of new facilities using rectangular distances. Oper. Res., Jan-Feb.

G.O. Wesolowsky, R.F. Love (1972) A nonlinear approximation method for solving a generalized rectangular distance Weber problem. Manag. Sci., 18, 56-63.

H.C. Wilson (1978) Least squares versus minimum absolute deviations estimation in linear models. Decision Sci., 322-335.

H.G. Wilson (1979) Upgrading transport costing methodology. Transportation J., 18, 49-55.

C.S. Withers (1986) The bias and skewness of L_1-estimates in regression. CSDA, 5, 301-303.

J.M. Wolfe (1979) On the convergence of an algorithm for a discrete L_p-approximation. Numer. Math., 32, 439-459.

R.S. Womersley (1986) Censored discrete linear L_1 approximation. SIAM J. Sci. Comput., 7, no.1, 105-122.

Y. Wu (1988) Strong consistency and exponential rate of the minimum L_1 norm estimatesK in linear regression models. CSDA 6(3), 285-296.

A. Wulff (1983) Numerische verfahren zur linearen orthogonalen L_p-regression. Diplomarbeit, Universitat Oldenburg.

J.D. Young (1971) Smoothing data with tolerances by use of linear programming. J. Inst. Math. Appl., 8, 69-79.

R. Zeckhauser, M. Thompson (1970) Linear regression with non-normal error terms. Rev. Econ. Stat., 52, 280-286.