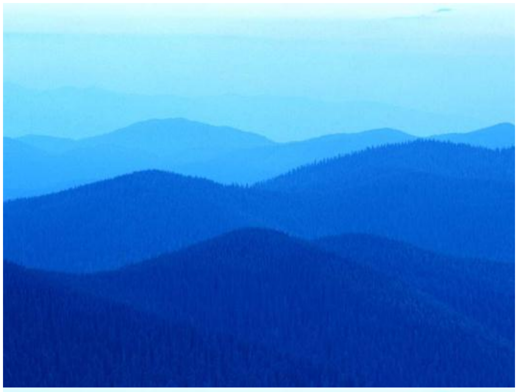
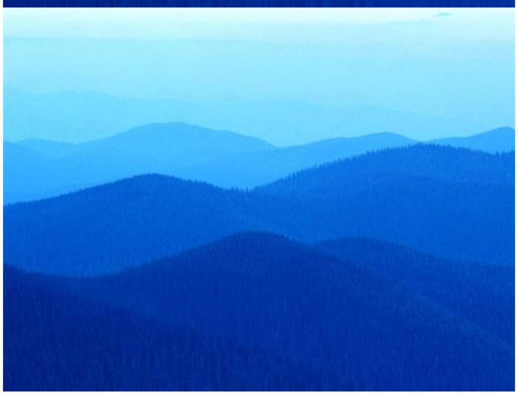
اقتصاد سنجی پیشرفته





دکتر حسین عباسی نژاد

دانشیار دانشکده اقتصاد دانشگاه تهران

فهرست مطالب

فهرست مطالب.. أ
پیش گفتار.. و
فصل اول.. ۱
۱- ۱ مدل رگرسیون خطی کلاسیک چند متغیره... ۱
۱-۱-۲ فرض‌های اولیه مدل کلاسیک... ۱
۱-۱-۳ روش حداقل مربعات معمولی (OLS)... ۲
۱-۱-۴ قضیه GAUSS – MARKOV (گوس و مارکف).. ۴
فصل دوم.. ۹
۲-۱ تجزیه تغییرات متغیر وابسته.. ۱۰
۲-۲ ضریب تعیین یا خوبی برازش: R^2.. ۱۳
فصل سوم.. ۱۹
۳-۱ تئوری مجانبی.. ۲۰
۳-۲ حد احتمالی دنباله تصادفی... ۲۰
۳-۳ قضیه خین چین.. ۲۲
۳-۴ قضایای اسلاتسکی.. ۲۲
۳-۵ سازگاری.. ۲۲
۳-۵-۱ سازگاری $\hat{\beta}_{OLS}$ تحت شرایط استاندارد مدل رگرسیون خطی........... ۲۳
۳-۵-۲ سازگاری $\hat{\sigma}^2$.. ۲۳
۳-۵-۳ توزیع احتمالی برای تخمین‌زننده‌های OLS... ۲۳
فصل چهارم... ۲۶
فصل چهارم.. ۲۷
۴-۱ معنی‌دار بودن ضرایب و فاصله اعتماد.. ۲۷
۵-۱ پیش‌بینی.. ۳۲
۶-۱ روش حداقل مربعات همراه با محدودیت.. ۳۶
۶-۲ استفاده از آزمون F برای بررسی قیدهای مدل رگرسیون خطی.............................. ۴۰
۶-۳ آزمون وجود رابطه بین متغیرهای توضیحی و وابسته... ۴۲

۶-۴ آزمون برابری دو مدل رگرسیون در دو وضعیت خاص اقتصادی مانند جنگ و صلح ۴۳
۶-۵ آزمون تغییرات ساختاری(Chow test (1960 .... ۴۴
۶-۵-۱ عرض از مبداهای مختلف ۴۷
۶-۷-۲ تغییردر برخی از ضرایب: ۴۹
۷-۱ مدل رگرسیون خطی تعمیم یافته ۵۲
۷-۲ روش حداقل مربعات تعمیم‌یافتـه (GLS) ۵۳
۷-۳ تئوری ایتکنز ۵۶
۸-۱ مدل رگرسیونی ظاهراً نامرتبط ۵۹
۹-۱ تخمین به روش حداکثر نمودن تابع درست‌نمایی ۶۴
۹-۲ مقایسه واریانس $\hat{\sigma}^2_{OLS}$ و $\tilde{\sigma}^2_{ML}$ ۶۷
۹-۳ ماتریس اطلاعات ۶۷
حد پایین رائو و کرامر ۶۸
۹-۳-۱ حد پایین رائو – کرامر ۶۹
۹-۴ استفاده از آزمونهای LM, LR, WALD ۷۰
۱۰-۱ واریانس ناهمسانی ۷۴
مدل ضرایب تصادفی ۷۷
۱۰-۲ واریانس ناهمسانی گروهی ۷۸
۱۰-۳ تبعات واریانس ناهمسانی درروش OLS برروی تخمین زننده‌ها ۸۰
۱۰-۴ انواع واریانس ناهمسانی ۸۰
واریانس ناهمسانی بصورت تابعی از متغیرهای مستقل یا برونزا ۸۰
۱۰-۵ -۱ آزمون تشخیص وجود واریانس ناهمسانی ۸۵
۱۰-۵-۲ آزمون گلدفلد - کوانت ۸۶
۱۰-۵-۲-۱ اشکالات وارد به روش آزمون فوق ۸۷
۱۰-۵-۳ آزمون بریوش - پایگان: (۱۹۷۹) ۸۷
۱۰-۵-۴ آزمون (۱۹۶۹) پارک - گلسجر ۸۸
۱۰-۵-۵ آزمون وایت ۸۹
۱۰-۵-۶ آزمون ضریب لاگرانژ برای واریانس ناهمسانی گروهی برای ۵ بنگاه با T مشاهده ۸۹
۱۱-۱ خود همبستگی ۹۲

۱۱-۲ تخمین تحت شرایط همبستگی پیاپی ۹۵
۱۱-۳ روش تخمین به روش حداکثر درست نمایی................................ ۹۶
۱۱-۴ مراحل روش تخمین کوکران اورکات:.. ۹۹
۱۱-۵ روش و مراحل کار Hilderth – LU.. ۹۹
۱۱-۶ روش و مراحل کار دوربین.. ۱۰۰
۱۱-۷ روش FGLS.. ۱۰۰
۱۱-۸ آزمون برای تشخیص همبستگی پیاپی (خود همبستگی)........................ ۱۰۰
۱۱-۸-۱ آزمون دوربین واتسون:.. ۱۰۱
۱۲-۱ هم خطی.. ۱۰۳
۱۲-۲ هم خطی کامل.. ۱۰۳
۱۲-۳ هم خطی ناقص.. ۱۰۴
۱۲-۴ تشخیص هم خطی.. ۱۰۶
۱۲-۴-۱ تجزیه واریانس ضرایب مدل رگرسیون.. ۱۰۷
فصل سیزدهم.. ۱۱۱
۱۳-۱ سیستم معادلات همزمان خطی.. ۱۱۱
۱۳-۲ روش متغیرهای ابزاری :.. ۱۱۳
۱۳-۳ مسئله تشخیص :.. ۱۱۴
۱۳-۴ تشخیص پارامترهای فرم ساختاری:.. ۱۱۵
۱۳-۵ نمادها ، فروض و قضایای حد مرکزی:.. ۱۱۸
۱۳- ۶ فروض مدل :.. ۱۲۰
۱۳-۷ حداقل مربعات غیرمستقیم:.. ۱۳۰
۱۳-۸ تخمین زننده با اطلاعات محدود:.. ۱۳۳
۱۳-۹ تخمین‌زننده با اطلاعات کامل :.. ۱۳۶
فصل چهاردهم.. ۱۳۹
تحلیل مدل سریهای زمانی.. ۱۳۹
۱- فراگرد تصادفی Stochastic Process.. ۱۳۹
$\{X_t : t \in T\}; T = N$.. ۱۳۹
۲- فراگرد تصادفی مانای ضعیفWeakly Stationary Stochastic Process...... ۱۳۹
۳- فراگرد تصادفی مانای قوی Strictly Stationary Stochastic Process ۱۴۰
۴- تابع اتوکو واریانس Auto Correlation Function(ACF) ۱۴۰

۵- تابع خود همبستگی Auto correlation Function **۱۴۰**
۶-تابع خود همبستگی جزیی Partial Autocorrelation Function **۱۴۰**
۷- White Noise **۱۴۱**
۸- مدلهای اتورگرسیو Autoregressive Model **۱۴۱**
۹- فراگرد تصادفی میانگین متحرک Moving Average Stochastic Process . **۱۴۲**
۱۰- فراگرد تصادفی Difference Stationary Process (DSP) **۱۴۲**
۱۱- گام تصادفی Random walk **۱۴۲**
۱۲- فراگرد تصادفی Trend Stationary process **۱۴۲**
۱۳- فراگرد نامانا و ریشه‌های واحد Non- Stationary Process and Unit Roots
.... **۱۴۳**
۱-۱۳- اهمیت ریشه‌های واحد ۱۴۳
۲-۱۳- مسئله آزمون فرضیه ۱۴۴
۱۴- آزمون ریشه واحد: Unit Root Test **۱۴۴**
۱۵- قدرت آزمون ریشه واحد **۱۴۵**
۱۶- مدلهای ARMA **۱۴۶**
۱-۱۶- AR(2) و AR(P) : مدل اتورگرسیو مرتبه دوم و مرتبه بالاتر P ۱۴۶
۲-۱۶- مرتبه هم جمعی Integration Order ۱۴۷
۳-۱۶- فراگرد میانگین متحرک مرتبه q ام، MA(q) ۱۴۸
۴-۱۶- مدلهای Autoregressive Moving Average Process ۱۴۸
۵-۱۶- فراگرد تصادفی ARIMA ۱۴۸
۶-۱۶- فراگرد تصادفی اتورگرسیو برداری Vector Auto regressive Process (VAR) ۱۴۹
رگرسیون هم جمعی The Co-Integration Regression **۱۴۹**
پیوست شماره ۱ .... **۱۵۸**
۱-۱ مفاهیم و نمادهای ابتدایی **۱۵۸**
۱.۲- خلاصه‌ای از عملیات پایه‌ای ماتریس **۱۵۹**
۱.۳- وابستگی خطی بردارها و مرتبه ماتریس **۱۶۰**
۱-۴- دترمینان ماتریس **۱۶۲**
۱-۵- معکوس ماتریس **۱۶۴**
۱-۶- ماتریس‌های افراز شده **۱۶۵**
۱-۷- اثر ماتریس **۱۶۶**

۸-۱- مقادیر ویژه (ریشه های ویژه) و بردارهای ویژه...... ۱۶۶
۹-۱- ماتریس‌های متعامد...... ۱۶۸
۱۰-۱- ماتریس‌های متقارن...... ۱۶۸
۱۱-۱- ماتریس هم قوه...... ۱۶۹
۱۲-۱- ماتریس های شبه معین و معین...... ۱۶۹
تمریـــــــن...... ۱۷۰
منابع:...... ۱۷۴

بنام آنکه فکرت آموخت

پیش گفتار

اقتصادسنجی یکی از شاخه‌های اقتصاد است که نظریه‌های اقتصادی، ریاضیات، آمار و علوم رایانه‌ای را برای مطالعهٔ پدیده‌های اقتصادی با یکدیگر جمع می‌کند. اقتصادسنجی هم یک گرایش در رشته اقتصاد است و هم یک ابزار بسیار قوی برای غالب اقتصاددانان و دانشمندان علوم اجتماعی که آن را برای مطالعهٔ علمی مسائل و موارد ویژه مورد استفاده قرار می‌دهند. هدف اولیهٔ اقتصادسنجی کاربردی نشان دادن نظریه‌های اقتصاد است. برای این منظور اقتصادسنجی شامل دامنهٔ وسیعی از فعالیت‌ها می‌باشد، که عبارتند از:

۱) فرمول دقیق ریاضی و نظریه‌های اقتصادی- اقتصاد ریاضی

۲) توسعه و گسترش تکنیک‌های آماری برای مدل اقتصادسنجی و آمار- نظریهٔ اقتصادسنجی

۳) توسعه روشها برای جمع‌آوری داده‌های اقتصادی- آمار اقتصاد

اقتصادسنجی به مقدار زیادی از آمار ریاضی، تجزیه و تحلیل عددی و علوم کامپیوتر استفاده می‌کند، با وجود این چیزی بیش از کاربرد مستقیم این تکنیک‌ها در مسائل اقتصادی است.

اقتصاددانان مانند سایر دانشمندان علوم اجتماعی نمی‌توانند شرایط را تحت کنترل خود در بیاورند و داده‌های مورد نیازشان را تولید کنند. عموماً ما در مرحلهٔ جمع‌آوری داده‌های آماری به‌صورت انفعالی عمل می‌کنیم و غالباً داده‌های آماری را مورد استفاده قرار می‌دهیم که کاملاً برای هدف‌هایی بدون ارتباط با تحقیق مورد نظر، تولید شده‌اند.

از طرف دیگر اقتصادسنجی‌دان به مجموعه‌ای غنی از نظریه‌های اقتصادی دسترسی دارد. تا اندازه‌ای نظریه‌های اقتصاد می‌تواند جایگزین طرح‌ها و آزمایش‌های تجربی قرارگیرد، یعنی در غالب اوقات، نظری، اطلاعاتی درباره انواع عوامل اقتصادی و متغیرهایی که می‌باید به هم مربوط باشند، ارائه می‌دهد. بعلاوه نظریهٔ اقتصادی می‌تواند اطلاعاتی را درباره اهمیت و آثار میان متغیرهای اقتصادی عرضه کند. در نتیجه وظیفه دیگر روش اقتصادسنجی این است که ابزاری را ارائه دهد که توسط آن اطلاعات غیرنمونه‌ای بتواند وارد تلاش‌های ساخت یک مدل شود. متاسفانه اقتصاد جزء علوم دقیقه نیست، درحالی که نظریه اغلب می‌تواند انواع کلی عواملی را که می‌باید در یک رابطهٔ اقتصادی درگیر شوند نشان دهد و در مورد فرم دقیق ریاضی نوعاً ساکت است. همچنین نمی‌تواند پیشنهادی راجع به چگونگی اندازه‌گیری توانایی فرد، درآمد دائمی و انتظارات ارائه

دهد. در حقیقت نظریه‌های اقتصادی ماهیتاً قیاسی می‌باشند. نتایج نظریه به‌طور منطقی از مفاهیم اولیه پیروی می‌کند. هیچ تضمینی وجود ندارد که داده‌های اقتصادی که مشاهده می‌کنیم به واسطهٔ فرایندی که توسط نظریهٔ خاص اقتصادی تشریح کرده‌ایم، تولید شده باشد.

همه این عوامل ما را متقاعدتر می‌کند که تمرین اقتصادسنجی توام با یک نااطمینانی است. یک راه مشخص و آشکار که ثابت کند نااطمینانی در داخل تجزیه و تحلیل‌های اقتصادی وجود دارد، واردکردن جملهٔ تصادفی در روابط اقتصادی می‌باشد که مدل‌های جبری را تبدیل به مدل‌های تصادفی می‌کند.

دلایلی به شرح زیر برای وارد کردن جزء اخلال در مدل وجود دارد.

اولاً: بحث می‌شود که مدل‌ها برای تحقیقات کاربردی است و تقریبی از روابط واقعی می‌باشد که داده‌های آماری را تولید می‌کند. نظریه‌ها، همان‌طور که مدل‌ها برای آزمون و یکپارچگی نتایج قضایا به‌کار برده می‌شود، تقریبی می‌باشند. دو نوع منبع و پایهٔ تقریب وجود دارد: یکی حذف عوامل از ملاحظات فکری که علی‌الاصول نامربوط است و دیگر این که تشخیص رابطهٔ تبعی ریاضی که در عمل انتخاب شده است و احتمالاً از یک فرم تبعی واقعی گرفته نشده است. جمله اختلال و تصادفی یا جملهٔ خطا بیانگر چنین صورتی از تقریب است.

ثانیاً: برای وارد کردن جملهٔ خطا آن است که حتی اگر تمام متغیرهای مربوط در یک مدل اقتصادسنجی داخل شوند و فرم تبعی نیز به‌طور صحیح طراحی شده بود، رابطهٔ بین متغیرها دقیق نیست، به‌دلیل این که در رفتار واحدهای اقتصادی جبر وجود ندارد.

در آخر، اقتصاددان‌ها اغلب متغیرهای اندازه‌گیری شده‌ای را به‌کار می‌برند که به درستی نشانگر عوامل نمی‌باشند که توسط نظریه بیان گردیده است. این اندازه‌گیری در خط به‌وسیلهٔ متغیرهایی که به‌عنوان نماینده به‌کار برده می‌شوند ایجاد شده و موجب می‌شوند که ارائه مدل‌های اقتصادسنجی از روابط اقتصادی، دقیق نباشند. البته این قضاوت‌ها مانعه‌الجمع نیستند و هرکدام یا همه آنها ممکن است به یک مدل خاص مربوط شود. وارد نمودن جمله تصادفی خطا در یک مدل اقتصادسنجی بدین معنی است که برای مشخص نمودن کامل یک مدل و رابطه نباید فقط مجموعه‌ای از متغیرها را که در روابط اقتصادی و فرم‌های ریاضی که بین متغیرها وجود دارند فرض نمود، بلکه می‌باید مجموعه‌ای از فروض را هم به روی توزیع جمله اخلال قرار داد.

وقتی مدل‌های اقتصادی‌سنجی به‌طور صحیح مشخص می‌شوند، نظریه‌های آماری، ابزار و وسایل تخمین‌های فاصله‌ای و نقطه‌ای پارامترها را فراهم می‌آورند، فرضیه‌ها و راه‌های ارزیابی نتایج و عملکرد آن قوانین را، آزمون می‌کنند.

اگر نااطمینانی منجر به این شود که اینها به یک مدل غیرصحیح ختم می‌شوند، آنگاه قوانین و قواعد و مراحل انجام شده ممکن است دیگر ویژگی‌ها و تبیین صحیح نداشته باشند، در نتیجه تمرکز دیگر نظریه اقتصادسنجی بر این است که به نتایج فروض غلط و توسعه متدهای مختلفی برای کشف و ارائه صحیح تصریح‌های مختلف منجر گردد. دلایلی که یک مدل اقتصادسنجی را می‌توان تخمین زد مختلف است:

اقتصاددان به‌طور طبیعی علاقه‌مند است ساختار اقتصاد و رفتار واحدهای اقتصادی را مورد مطالعه قرار دهد. امنیت‌های اقتصادی در روابط اقتصادی اقتصاددان را وامی‌دارد که صحیح بودن نظریه‌های اقتصادی را آزمون کند و اگر مدل تخمین‌زده شده غیرمفید و نامناسب تشخیص داده نشد، اطلاعات تعدادی برای پدیده‌های اقتصادی به‌دست می‌دهد. این اطلاعات تعدادی علاوه بر این که دانش عمومی ما را افزایش داده‌اند، چند وظیفه دیگر از آن می‌توان استنباط کرد.

روابط اقتصادی تخمین زده شده پایه و اساس پیش بینی متغیرهای مهم اقتصادی، قرار می‌گیرد. این پیش‌بینی‌ها ممکن است برای سیاستگذاران اقتصادی ارزشمند باشد. کسانی که دوست دارند آثار و حساسیت‌هایی را که در تغییر متغیرهای اقتصادی تحت کنترل آنها به‌وجود می‌آورد را پیش بینی کنند.

فصل اول

۱

فصل اول

۱- ۱ مدل رگرسیون خطی کلاسیک چند متغیره

مدل خطی زیر را در نظر بگیرید:

$$Y_t = \beta_1 X_{t1} + \beta_2 X_{t2} + \ldots + \beta_k X_{tk} + \varepsilon_t \qquad t = 1, \ldots, T$$

Y_tعبارت است از مشاهده t ام از متغیر وابسته، X_{ti} عبارت است از مشاهده t ام بر روی متغیر مستقل i ام. β_i ضریب متناظر با متغیر مستقل i در مدل رگرسیون می‌باشد. و بالاخره ε_t جمله خطا مربوط به مشاهده t ام می‌باشد. متغیر مستقل را می‌توانیم رگرسور و جمله خطا را متغیر بر هم‌زننده نیز نام‌گذاری کنیم. T عبارتست از تعداد مشاهدات و K تعداد رگرسورها یا متغیرهای توضیحی در مدل میباشند.

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1k} \\ x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ x_{T1} & x_{T2} & \cdots & x_{Tk} \end{bmatrix}$$

مدل فوق را می‌توانیم بصورت زیر بنویسیم:

$$Y = X\beta + \varepsilon$$

که Y بردار $T \times 1$ ، X ماتریس $T \times K$ و ε بردار $T \times 1$ می‌باشد بنابراین:

$$Y' = (Y_1, Y_2, \cdots, Y_T) \qquad \varepsilon' = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_T) \qquad \beta' = (\beta_1, \ldots, \beta_k)$$

۱-۱-۲ فرض‌های اولیه مدل کلاسیک

۱- X ماتریسی غیرتصادفی[1] با مرتبه[2] $K \le T$ می‌باشد بطوریکه در نمونه‌های بزرگ که تعداد مشاهدات به سمت بی‌نهایت میل می‌کند:

$$Lim_{T \to \infty}(X'X / T) = Q$$

Q یک ماتریس غیرمنفرد (معکوس‌پذیر) است.

---

1. Nonstochastic
2. Rank

۲- بردار ε شامل خطاهای غیرقابل مشاهده است به نحویکه شرایط زیر را برآورده سازد:

$$i)E(\varepsilon)=0$$

$$ii)E(\varepsilon\varepsilon')=\sigma^2 I$$

باتوجه به شرط (i) می‌توانیم میانگین شرطی متغیر وابسته Y را بصورت زیر تعریف کنیم:

$$E(y|x)=X\beta$$

مفهوم عبارت بالا این است که، امید ریاضی متغیرهای تصادفی y_t توابعی خطی از متغیرهای مستقل X_{ti} می‌باشند. بدلیل اینکه متغیر y_t توسط متغیرهای X_{ti} توضیح داده می‌شود، آن را متغیر وابسته می‌نامیم.

توجه: فرض اول کلاسیک رفتار متغیرهای مستقل را محدود می‌کند بطوریکه مثلاً تعداد آنها نمی‌تواند با سرعت افزایش مشاهدات افزایش یابد، هم‌چنین هیچ یک از این متغیرها نمی‌تواند روی همه مشاهدات صفر باشد و یا رابطه خطی با متغیرهای دیگر مستقل داشته باشد.

مثال (۱)

$$Y_t=X_{t1}\beta_1+X_{t2}\beta_2+\varepsilon_t$$

فرض کنید $X_{t2}=CX_{t1}$ در اینصورت نشان دهید که :

$$Lim_{T\to\infty}\left(\frac{X'X}{T}\right)$$

یک ماتریس منفرد است.

فرض(۱) هم‌چنین وجود هم‌خطی بین متغیرهای مستقل را نفی کرده است و آن بخاطر مرتبه تعریف شده X است. چون هیچ ستونی از ماتریس X نمی‌تواند ترکیب خطی از سایر ستونهای X باشد.

فرض(۲) از فروض کلاسیک توزیع احتمال جمله اخلال را مشخص می‌کند بطوریکه این جملات همگی بطور مستقل و یکنواخت با میانگین صفر و واریانس ثابت σ^2 ، توزیع شده‌اند.

۱-۱-۳ روش حداقل مربعات معمولی[۳] (OLS)

مسئله تخمین مدل خطی کلاسیک مستلزم تخمین پارامترهای مجهول $\beta_1,\ldots,\beta_k$ و σ^2 می‌باشد. روش حداقل مربعات معمولی، مقادیر $\beta_k,\ldots,\beta_2,\beta_1$ را بنحوی انتخاب می‌کند که مجموع مربعات خطاها مینیمم شود، یعنی :

$$S(\hat{\beta})=\sum_{t=1}^{T}\left(Y_t-\beta_1X_{t1}-\ldots-\beta_kX_{tk}\right)^2$$

3 . Ordinary Least Square

که بصورت ماتریس خواهیم داشت:

$$S = (Y - X\beta)'(Y - X\beta)$$

شرایط لازم برای مینیمم کردن تابع بالا با مشتق‌گیری نسبت به β و سـپس برابـر صـفر قـرار دادن آن بدست خواهد آمد، لذا خواهیم داشت:

$$S = Y'Y - \beta'X'Y - Y'X\beta + \beta'X'X\beta$$
$$S = Y'Y - 2\beta'X'Y + \beta'X'X\beta$$

معادلات نرمـــال

یـــا :

$$\frac{\partial S}{\partial \beta} = -2X'Y + 2X'X\hat{\beta} = 0$$

$$X'Y = X'X\hat{\beta} \Rightarrow \hat{\beta} = (X'X)^{-1}X'Y$$

ماتریس مشتق دوم این تابع عبارتســت از :

$$\frac{\partial^2 S}{\partial\beta\partial\beta'} = 2X'X$$

این ماتریس ، مثبت معین[۴] است، لـذا $\hat{\beta}$ تابع S را مینیمم می‌کند.

ماتریس واریانس- کوواریانس $\hat{\beta}$ نیز یه صورت زیر است:

$$VC(\hat{\beta}) = E(\hat{\beta} - E(\hat{\beta}))(\hat{\beta} - E(\hat{\beta}))' = E(\hat{\beta} - \beta)(\hat{\beta} - \beta)'$$
$$= E(X'X)^{-1}X'\varepsilon\varepsilon'X(X'X)^{-1}$$
$$= (X'X)^{-1}X'E(\varepsilon\varepsilon')X(X'X)^{-1}$$

چون $E(\varepsilon\varepsilon') = \sigma^2 I$ می‌باشد و امید ریاضی عملگری است که از کمیت‌های غیرتصادفی عبـور نمیکنـد لذا ماتریس واریانس – کوواریانس $\hat{\beta}$ برابر است با :

$$VC(\hat{\beta}) = \sigma^2 (X'X)^{-1}$$

4. (Positive Definite)

۴-۱-۱ قضیه GAUSS – MARKOV (گوس و مارکف)

این قضیه می‌گوید تخمین‌زننده OLS بهترین تخمین‌زننده خطی و بدون تورش[۵] در میان تخمین‌زننده‌های خطی و بدون تورش است که به آن $BLUE$ [۶] گفته می‌شود.

$$\hat{\beta} = (X'X)^{-1} X'Y$$

$$\hat{\beta} = (X'X)^{-1} X'(X\beta + \varepsilon)$$
$$= (X'X)^{-1} X'X\beta + (X'X)^{-1} X'\varepsilon$$

$$\hat{\beta} = \beta + (X'X)^{-1} X'\varepsilon$$

چون X غیرتصادفی است.

$$E\left(\hat{\beta}\right) = \beta + (X'X)^{-1} X'E(\varepsilon)$$

چون $E(\varepsilon) = 0$ است.

$$\Rightarrow E\left(\hat{\beta}\right) = \beta$$

و لـــذا $\hat{\beta}$ بدون تورش[۷] است.

برای نشان دادن کارآیی (حداقل واریانس داشتن) تخمین‌زننده $\hat{\beta}$ و اینکه OLS بهترین است، اجازه دهید تخمین‌زننده خطی دیگری را در نظر بگیریم. ابتدا ناتور بودن را بررسی می‌کنیم:

$$\beta^* = HY$$

بطوریکه:

$$H = (X'X)^{-1} X' + C$$

که یک ماتریس ثابت است، سپس:

$$\beta^* = HX\beta + H\varepsilon \quad , \quad E(\beta^*) = E(HX\beta)$$

شرط بدون تورش بودن ایجاب می‌کند که : $HX = I$ باشد.

$$HX = (X'X)^{-1} X'X + CX = I + CX$$

لذا در صورتی $HX = I$ خواهد بود که $CX = 0$ و c غیرتصادفی باشد که در بالا ثابت در نظر گرفتیم. پس تخمین‌زننده β^* خطی و بدون تورش است.

اما ماتریس واریانس کوواریانس β^* عبارتست از:

$$E(\beta^* - \beta)(\beta^* - \beta)' = E(H\varepsilon\varepsilon'H')$$
$$= E\left[(X'X)^{-1}X' + C\right]\varepsilon\varepsilon'\left[X(X'X)^{-1} + C'\right]$$
$$= \sigma^2\left[(X'X)^{-1}X'X(X'X)^{-1}X'C' + CX(X'X)^{-1} + CC'\right]$$
$$= \sigma^2 (X'X)^{-1} + \sigma^2 CC'$$

چون $CX = 0$ است.

---

5. Bias
6 . Best Linear Unbiased Estimator
7. Unbiased

ولی CC' یک ماتریس مثبت شبه معین است. لذا واریانس β^*، بزرگتر از واریانس $\hat{\beta}$ است.

چون اختلاف $\beta^* - \hat{\beta}$ یک ماتریس مثبت شبه‌معین است بنابراین تخمین‌زن $\hat{\beta}$ کوچکترین واریانس را دارد و در نتیجه نسبت به سایر تخمین‌زننده‌های خطی کارآ است و چون یک تخمین‌زننده بدون تورش نیز است به همین دلیل OLS را یک تخمین‌زننده $BLUE$ (بهترین، خطی، بدون تورش) می‌نامند.

یادآوری:

اولاً : قضیه گوس- مارکف تنها در مورد نمونه‌های کوچک صادق است. یعنی باتوجه به فروض مدل کلاسیک نتیجه قضیه برای هر اندازه‌ای از نمونه که $(T \geq K)$ باشد تخمین زننده بدون تورش است.

ثانیاً : ما در اینجا نوع توزیع جمله اخلال را مشخص نکردیم پس ضرورتاً نبایستی ε دارای توزیعی نرمال باشد.

۱-۱-۵ قضیه: در مدل رگرسیون خطی کلاسیک، تخمین σ^2 یک تخمین بدون تورش است که برابر است با :

$$\hat{\sigma}_3^2 = \frac{\hat{\varepsilon}'\hat{\varepsilon}}{T-K}$$

بطوریکه:

$$\hat{\varepsilon} = Y - X\hat{\beta}$$

اثبات: بدون تورش بدون $\hat{\sigma}^2$ با گرفتن امید ریاضی از مربع تخمین جملات اخلال بدست می‌آید.

$$\hat{\varepsilon} = Y - X\hat{\beta} = Y - X(X'X)^{-1}X'Y = \left[I - X(X'X)^{-1}X'\right]Y$$
$$= MY \qquad \text{که} \quad M = I - X(X'X)^{-1}X'$$

یا:

$$\hat{\varepsilon} = \left[I - X(X'X)^{-1}X'\right] \quad [X\beta + \varepsilon]$$
$$= X\beta - X(X'X)^{-1}X'X\beta + M\varepsilon = M\varepsilon$$

توجه: ماتریس M یک ماتریس متقارن و همقوه[8] است یعنی:

$$M' = M$$
$$MM = M^2 = M$$

به نحویکه مرتبه M با اثر[9]، M برابر است.

بنابراین مجموع مربعات $\hat{\varepsilon}$ عبارتست از:

---

8 . Idempotent
9 . Trace

$$E\left(\hat{\varepsilon}'\hat{\varepsilon}\right)=E\left[\varepsilon'M'M\varepsilon\right]=E\left(\varepsilon'M\varepsilon\right)$$

چون $\varepsilon'M\varepsilon$ یک اسکالر است و اثر آن با خودش برابر است لـذا :

$$E\left(\varepsilon'M\varepsilon\right)=Etr\left(\varepsilon'M\varepsilon\right)$$

$$=Etr\left(M\varepsilon\varepsilon'\right)$$ چون $tr\left(AB\right)=tr\left(BA\right)$

$$=trE\left(M\varepsilon\varepsilon'\right)$$ چون tr یک تابع خطی است.

$$=trM.E\left(\varepsilon\varepsilon'\right)$$ چون M غیرتصادفی است.

$$=trM\sigma^2I=\sigma^2trM$$

داشتیم:

$$M=I-X\left(X'X\right)^{-1}X'$$

که:

$$trM=trI_t-tr\left(X\left(X'X\right)^{-1}X^{-1}X'\right)$$

چون $tr\left(A+B\right)=trA+trB$

$$=trI_T-tr\left(X'X\left(X'X\right)^{-1}\right)$$ چون $trAB=trBA$

$$=trI_T-trI_K=T-K$$

لـــذا :

$$E\left(\hat{\varepsilon}'\hat{\varepsilon}\right)=\sigma^2\left(T-K\right)$$

بنابراین:

$$E\left(\hat{\varepsilon}'\hat{\varepsilon}/T-K\right)=\sigma^2$$

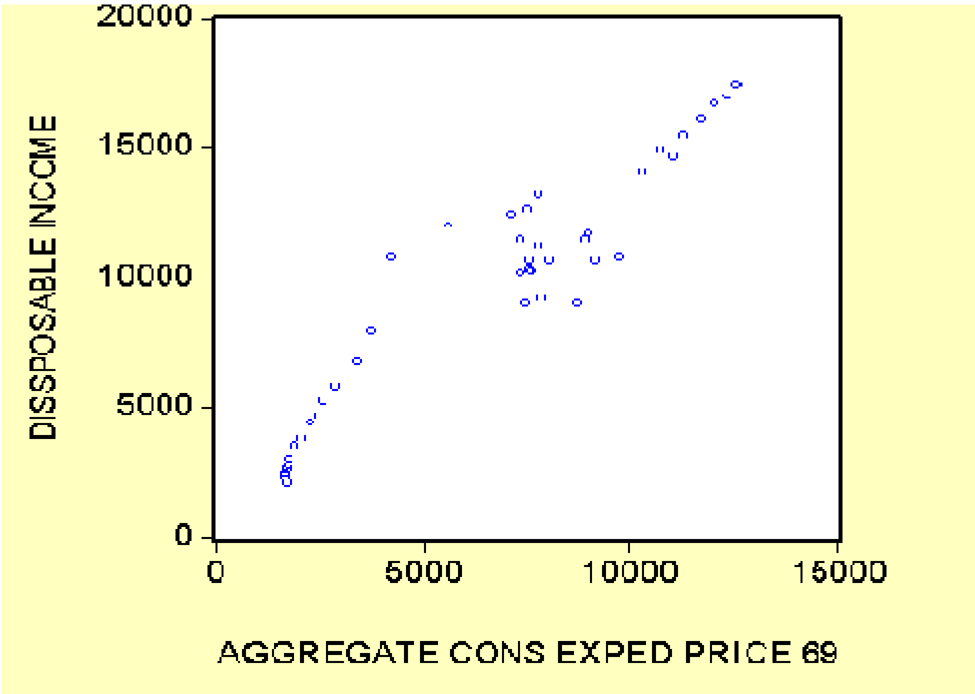
بدون تورش بودن $\hat{\sigma}^2$ ثابت شد.

مثال:

در شکل ۱ پراکندگی داده‌های آماری مربوط به درآمد قابل تصرف و مخارج بخش خصوصی آمده است.

شکل ۱

پراکندگی داده های مخارج بخش خصوصی در مقابل درآمد قابل تصرف ۱۳۷۸-۱۳۳۸



جدول ۱

نتایج تخمین تابع مصرف

Dependent Variable: PCE
Method: Least Squares
Date: 02/14/03 Time: 05:22
Sample: 1338 1378
Included observations: 41

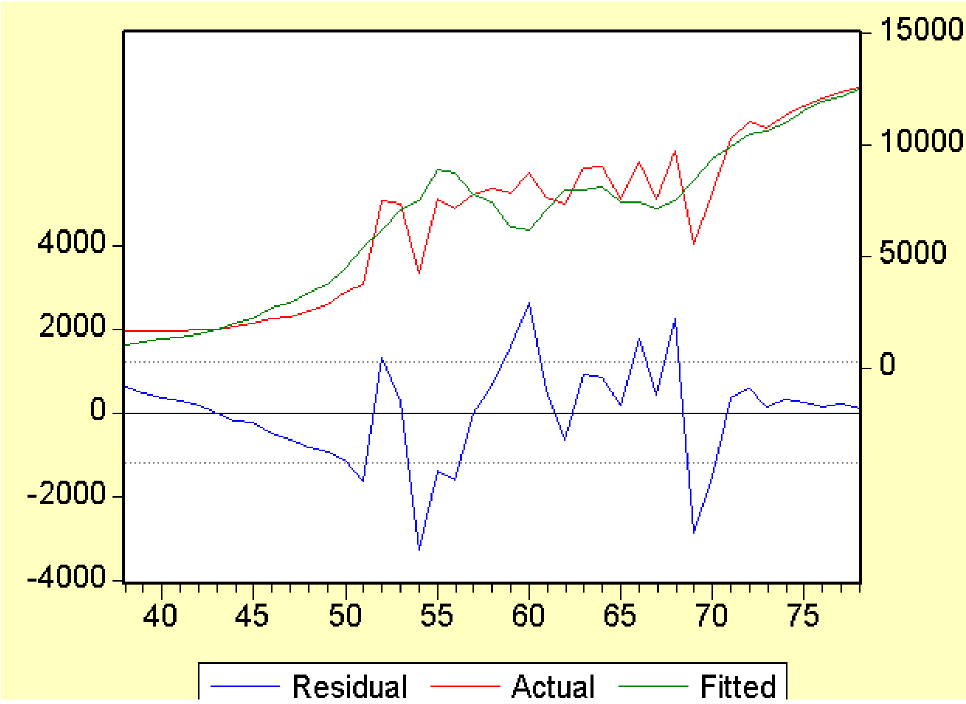
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-599.2746	436.9501	-1.371494	0.1781
PDI	0.750105	0.041369	18.13194	0.0000

R-squared	0.893955	Mean dependent var	6564.613
Adjusted R-squared	0.891236	S.D. dependent var	3623.126
S.E. of regression	1194.887	Akaike info criterion	17.05704
Sum squared resid	55682458	Schwarz criterion	17.14063
Log likelihood	-347.6694	F-statistic	328.7674
Durbin-Watson stat	1.434111	Prob(F-statistic)	0.000000

۱۳

شکل ۲

ترسیم متغیر مشاهده شده- تخمین آن و جملات پسماند



فصل دوم

فصل دوم

فصل دوم

۲-۱ تجزیه تغییرات متغیر وابسته

بطور کلی در یک مدل رگرسیون ، کل تغییرات متغیر وابسته برابر است با مجموع تغییرات توضیح داده شده و مجموع خطاها که معمولا بصورت زیر نوشته میشود:[۱۰]

$$TSS= ESS + RSS \qquad (۱)$$

مجموع خطاها + مجموع تغییرات توضیح داده شده =کل تغییرات

و یا میتوانیم آنرا به صورت زیر بنویسیم:

$$\sum(y_t - \bar{y})^2 = \sum(\hat{y}_t - \bar{\hat{y}})^2 + \sum \hat{u}_t^2$$

برای اثبات اینکه رابطه فوق برقرار است ابتدا قضایای زیر را بر رسی می‌کنیم.

---

10 . TSS = Total Sum of Square
ESS = Explained Sum of Square
RSS = Residual Sum of Square

قضیه ۱- می دانیم که مشاهده t ام از متغیر وابسته برابر است با برآورد همان مشاهده بعلاوه برآورد خطای متناظر با مشاهده t ام از متغییر وابسته یعنی:

$$y_t = \hat{y}_t + \hat{u}_t \qquad (۲)$$

حال چنانچه بردار مشاهدات را برای y_t ، $\hat{y}_t$ و $\hat{u}_t$ درنظر بگیریم، مجموع مجذورات طرفین رابطه (۲) با استفاده از جبر خطی بصورت زیر خواهد شد:

$$y'y = (\hat{y} + \hat{u})'(\hat{y} + \hat{u})$$
$$y'y = \hat{y}'\hat{y} + \hat{u}'\hat{y} + \hat{y}'\hat{u} + \hat{u}'\hat{u}$$

می توان نشان داد که جملات دوم و سوم در سمت راست برابر صفر است، چون داریم:

$$\begin{aligned}\hat{y}'\hat{u} &= [x\hat{\beta}]'\hat{u}\\ &= [x(x'x)^{-1}x'y]'\hat{u}\\ &= y'x(x'x)^{-1}x'\hat{u}\end{aligned}$$

از طرفی به دلیل عدم ارتباط بین جملات پسماند و متغیرهای مستقل $x'\hat{u} = 0$ است

چون: $x'\hat{u} = x'My = x'(I - x(x'x)^{-1}x')y = 0$

پس ثابت شد که:

$$\hat{y}'\hat{u} = \hat{u}'\hat{y} = 0$$

لذا داریم: (۳)

$$y'y = \hat{y}'\hat{y} + \hat{u}'\hat{u}$$

قضیه ۲- چنانچه مدل رگرسیون خطی دارای عرض از مبدا باشد میانگین جملات پسماند برابر با صفر است. یعنی : $\bar{\hat{u}} = 0$

اثبات:

از قبل داشتیم $x'\hat{u} = 0$ چنانچه آنرا به صورت ماتریسی باز بنویسیم داریم:

$$x'\hat{u} = [x_{.1,\dots,}\ x_{.k}]'\hat{u} = \begin{bmatrix} x'_{.1}\hat{u} \\ \vdots \\ x'_{.k}\hat{u} \end{bmatrix}$$

$$= \begin{bmatrix} \sum x_{t1} \quad \hat{u}_t \\ \vdots \\ \sum x_{tk}\hat{u}_t \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

رابطه فوق زمانی برقرار است که تک‌تک عناصر بردار اول برابر صفر باشد. لذا جمله اول این بردار برابر صفر است. پس داریم:

$$\sum x_{t1}\hat{u}_t = 0$$

اگر مدل دارای عرض از مبدا باشد $x_{t1} = 1$ است و خواهیم داشت:

$$\sum \hat{u}_t = 0 \Rightarrow \bar{\hat{u}} = \frac{1}{T} \sum \hat{u}_t = 0$$

قضیه ۳- اگر مدل رگرسیون خطی همچنان دارای عرض از مبدا باشد، میانگین مشاهدات بر روی متغیرهای وابسته برابر است با میانگین برآورد آنها، یعنی: $\bar{\hat{y}} = \bar{y}$

اثبات: از رابطه (۲) داریم:

$$\sum y_t = \sum \hat{y}_t + \sum \hat{u}_t$$
$$\frac{1}{T}\sum y_t = \frac{1}{T}\sum \hat{y}_t + \frac{1}{T}\sum \hat{u}_t$$

چون جمله دوم سمت راست برابر صفر است، پس :

$$\Rightarrow \bar{\hat{y}} = \bar{y}$$

قضیه ۴- حال میتوان با توجه به قضایای (۱)، (۲)، (۳) رابطه (۱) را اثبات نمود:

چون $\bar{\hat{y}} = \bar{y}$ می باشد از سمت راست رابطه(۳) مقدار $T\,\bar{\hat{y}}^2$ و از سمت چپ آن مقدار $T\,\bar{y}^2$ را کسر می کنیم:

و لذا حکم ثابت شد:

$$y'y - T\bar{y}^2 = \hat{y}'\hat{y} - T\bar{\hat{y}}^2 + \hat{u}'\hat{u}$$

$$\sum(y_t - \bar{y})^2 = \sum(\hat{y}_t - \bar{\hat{y}})^2 + \sum \hat{u}_t^2$$

چون:

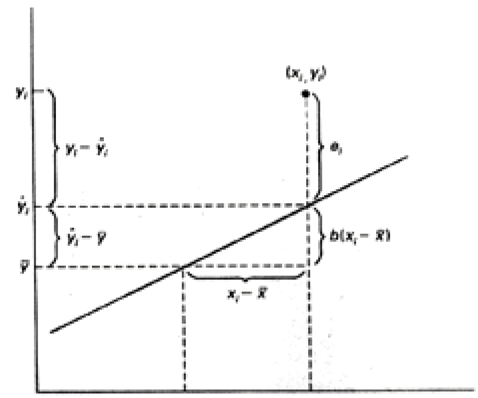
$$\sum(y_t - \bar{y})^2 = \sum y_t^2 - T\bar{y}^2$$

$$\sum(\hat{y}_t - \bar{\hat{y}})^2 = \sum \hat{y}_t^2 - T\bar{\hat{y}}^2$$

پس چنانچه مدل دارای عرض از مبدا باشد، مجموع تغییرات متغییر وابسته برابر است با مجموع تغییرات توضیح داده شده بعلاوه مجموع خطاها که آنرا می توان بصورت فوق نوشت.

شکل ۱

تجزیه تغییرات متغیر وابسته



۲-۲ ضریب تعیین یا خوبی برازش: R^2

R^2 یا ضریب تعیین، قدرت توضیح دهندگی مدل را نشان می‌دهد. یعنی نشان می‌دهد که چند درصد تغییرات متغیر وابسته توسط متغیرهای توضیحی، توضیح داده شده‌اند و عبارتست از نسبت تغییرات توضیح داده شده به کل تغییرات:

$$R^2 = \frac{ESS}{TSS} = \frac{TSS - RSS}{TSS} = 1 - \frac{RSS}{TSS}$$

بعبارت دیگر R^2 عبارتست از :

$$R^2=\frac{\sum(\hat{y}_t-\bar{\hat{y}}_t)^2}{\sum(y_t-\bar{y})^2}=1-\frac{\sum \hat{u}_t^{\,2}}{\sum(y_t-\bar{y})^2}$$

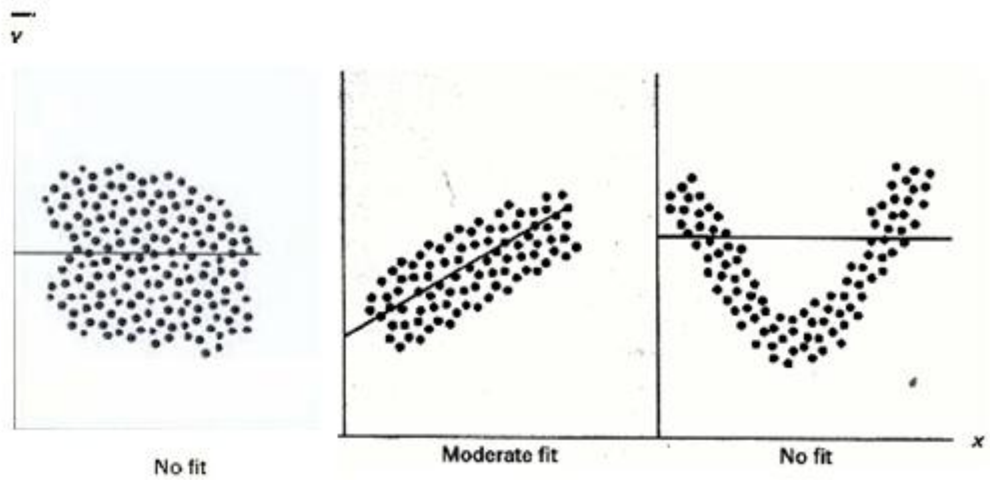
چنانچه مدل دارای عرض از مبدا باشد $0\leq R^2\leq 1$ می باشد.

صورت دیگری میتوان برای R^2 در نظر گرفت:

$$R^2 \;=\; \frac{[\frac{1}{T}\sum(y_t-\bar{y})(\hat{y}_t-\bar{\hat{y}})\;]^2}{\frac{1}{T}\sum(y_t-\bar{y})^2\,\frac{1}{T}\sum(\hat{y}_t-\bar{\hat{y}})^2}$$

رابطه بالا نشان میدهد که R^2 عبارت است از ضریب همبستگی ساده بین متغییر وابسته و برآورد آن یا بعبارت دیگر نشان دهنده ضریب همبستگی بین مقدار واقعی متغیر وابسته و مقدار برازش شده آن میباشد.

توجه: چنانچه مدل دارای عرض از مبدا نباشد مقدار ضرورتا بین صفر و یک نخواهد بود ممکن است مقدار منفی نیز اختیار کند.



مثال: با توجه به داده های آماری زیر مدل اثرات درآمد ملی و نرخ بهره را برروی سرمایه گذاری بررسی میکنیم.

YEAR	GNP	INVEST	CPI	INTEREST
1968	873.4	133.3	82.54	5.16
1969	944	149.3	86.79	5.87
1970	992.7	144.2	91.45	5.95
1971	1077.6	166.4	96.01	4.88
1972	1185.9	195	100	4.5
1973	1326.4	229.8	105.75	6.44
1974	1434.2	228.7	115.08	7.83
1975	1549.2	206.1	125.79	6.25
1976	1718	257.9	132.34	5.5
1977	1918.3	324.1	140.05	5.46
1978	2163.9	386.6	150.42	7.46
1979	2417.8	423	163.42	10.28
1980	2633.1	402.3	178.64	11.77
1981	2937.7	471.5	195.51	13.42
1982	3057.5	421.9	207.23	11.02

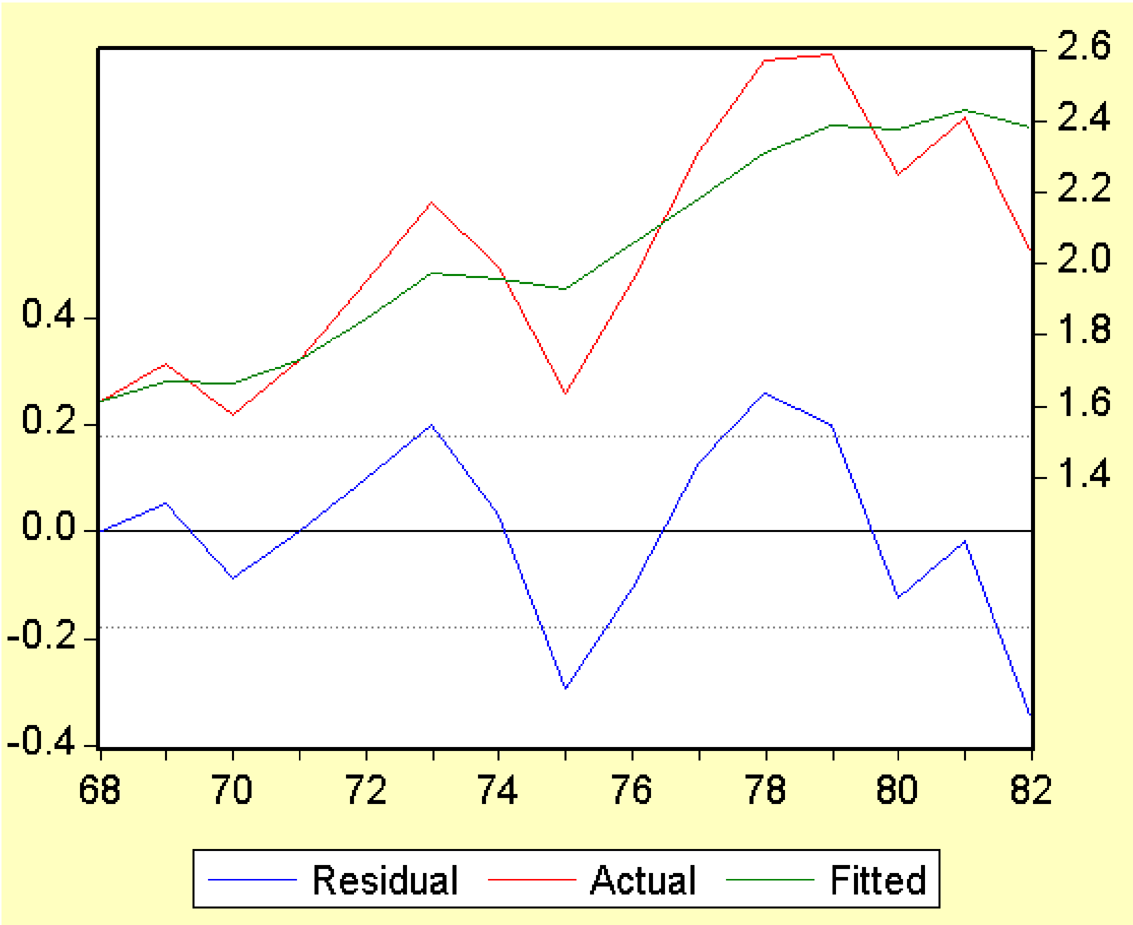
مدل اول رابطه بین سرمایه گذاری حقیقی و تولید ناخالص ملی حقیقی را بررسی میکند که نتایج در جدول ۱ آمده است.

جدول ۱

Dependent Variable: RINVEST
Method: Least Squares
Date: 10/31/04 Time: 13:50
Sample: 1968 1982
Included observations: 15

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.329482	0.386854	-0.851696	0.4098
RGNP	0.183624	0.029836	6.154420	0.0000

R-squared	0.744481	Mean dependent var	2.034337
Adjusted R-squared	0.724826	S.D. dependent var	0.341145
S.E. of regression	0.178955	Akaike info criterion	-0.479804
Sum squared resid	0.416321	Schwarz criterion	-0.385398
Log likelihood	5.598533	F-statistic	37.87689
Durbin-Watson stat	1.240725	Prob(F-statistic)	0.000035



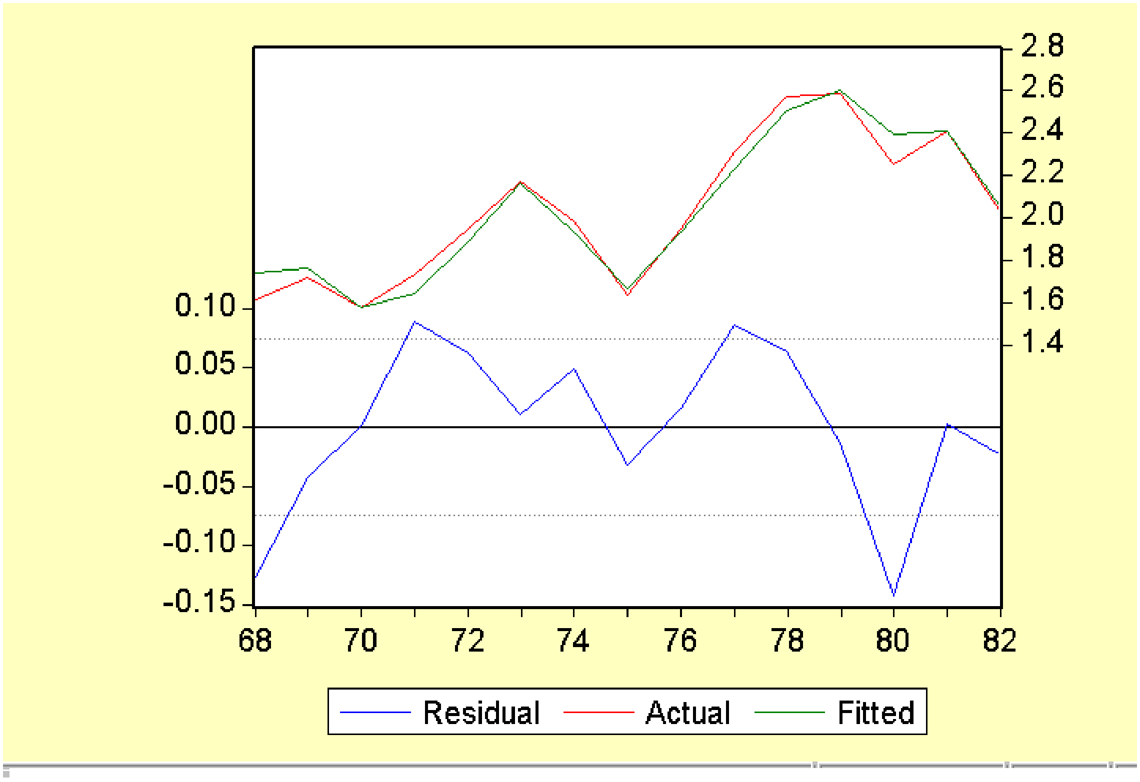
در مدل بعدی به منظور روند زدایی از متغیرها متغیر زمانرا نیز وارد میکنیم. نتایج در جدول ۲ آمده است و ملاحظه میشود که ضریب سرمایه‌گذاری تغییر محسوسی نموده است و این نشانگر آنست که اثرات متغیر زمان نه تنها باعث کاهش اثرات تولید ناخالص ملی حقیقی نگردیده است بلکه موجب افزایش آن نیز گردیده است. چون گمان میرفت وجود مروند در متغیرهای سری زمانی منجر به ایجاد رابطه کاذب بین متغیرهای نامربوط میشود. در حالیکه ملاحظه میکنیم در اینجا روند زمانی حتی اثرات رو به پایین بر روی سریهای زمانی داشته است.

جدول ۲

Dependent Variable: RINVEST
Method: Least Squares
Date: 11/02/04 Time: 10:13
Sample: 1968 1982
Included observations: 15

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	333.5816	41.58081	8.022489	0.0000
RGNP	0.653577	0.059800	10.92940	0.0000
YEAR	-0.172132	0.021435	-8.030472	0.0000

R-squared	0.959913	Mean dependent var	2.034337
Adjusted R-squared	0.953231	S.D. dependent var	0.341145
S.E. of regression	0.073776	Akaike info criterion	-2.198705
Sum squared resid	0.065315	Schwarz criterion	-2.057094
Log likelihood	19.49028	F-statistic	143.6729
Durbin-Watson stat	1.218399	Prob(F-statistic)	0.000000



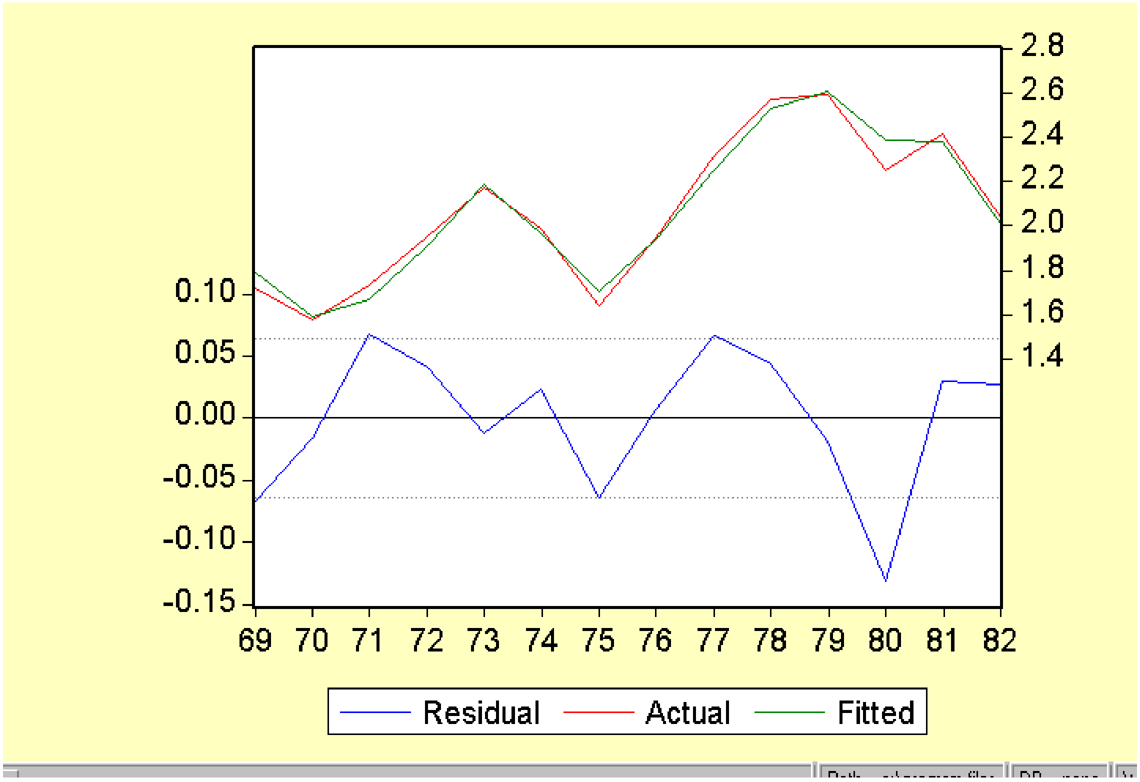
حال اجازه دهید به منظور تصریح بهتر مدل متغیر نرخ بهره حقیقی را که حاصل تفریق نرخ تورم از نرخ بهـره اسمی است را در مدل وارد کنیم. نتایج در جدول ۳ آمده است.

جدول ۳

Dependent Variable: RINVEST
Method: Least Squares
Date: 11/02/04 Time: 10:40
Sample(adjusted): 1969 1982
Included observations: 14 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	350.2740	37.69605	9.292061	0.0000
RGNP	0.672606	0.052630	12.78000	0.0000
YEAR	-0.180699	0.019417	-9.306406	0.0000
RINTEREST	-0.008504	0.010865	-0.782703	0.4519

R-squared	0.971428	Mean dependent var	2.064292
Adjusted R-squared	0.962857	S.D. dependent var	0.332923
S.E. of regression	0.064163	Akaike info criterion	-2.419829
Sum squared resid	0.041169	Schwarz criterion	-2.237242
Log likelihood	20.93881	F-statistic	113.3323
Durbin-Watson stat	1.776977	Prob(F-statistic)	0.000000



فصل سوم

فصل سوم

۳-۱ تئوری مجانبی[11]

به‌طور کلی تخمین زننده $\hat{\theta}_T$ تابعی از متغیرهای تصادفی Y_t است یعنی:

$$\hat{\theta}_T = F(Y_1, Y_2, ..., Y_T)$$

چنانچه بتوانیم توزیع $\hat{\theta}_T$ را از توزیع Y_t ها محاسبه کنیم در آنصورت قادریم ویژگی‌های $\hat{\theta}_T$ را از طریق توزیع $\hat{\theta}_T$، تجزیه و تحلیل کنیم. اگر همه اطلاعات لازم را درباره توزیع نداریم ممکن است به دو گشتاور اولیه (میانگین و واریانس) اکتفا کنیم. چنانکه آن هم میسر نباشد ویژگی‌های $\hat{\theta}_T$ را می‌توانیم وقتی حجم نمونه بزرگ می‌شود تجزیه و تحلیل کنیم این عمل را تئوری مجانبی می‌گویند.

۳-۲ حد احتمالی دنباله تصادفی[12]

دنبال تصادفی $Y_1, Y_2, ..., Y_T$ را که دارای توزیع $F_1(.), F_2(.)....F_n(.)$ می‌باشد به‌طور احتمالی به‌سمت ثابت C گرایش پیدا می‌کند اگر:

$$\mathrm{Lim}\, p\left[\left|Y_t - c\right| > \varepsilon\right] = 0 \qquad \text{for} \qquad \varepsilon > 0$$

نامساوی داخل پرانتز یا صحیح است یا غلط، احتمال درست بودن آن توسط تابع توزیع $F_n(.)$ مربوط به ε, C, Y_T تعیین می‌شود.

دنباله تصادلی متغیرهای $Y_1, Y_2, ..$ در احتمال به‌سمت ثابت C گرایش پیدا می‌کند اگر حد این دنباله احتمالات برای $\varepsilon > 0$ صفر شود.

$$P \lim_{T \to \infty} Y_t = C$$

---

11 - Asymptotic Theory.

12 - Convergence in Probability of a Random Sequence.

یعنی حد احتمالی دنباله تصادفی برابر با C می‌باشد این مفهوم دقیقاً برابر است با اینکه بگوییم: دنباله $Y_1,...,Y_T$ در احتمال به سمت C گرایش پیدا می‌کند، اما کلمه حد احتمال دارای این امتیاز است که روی تفاوت آن با حد معمولی تاکید می‌کند وقتی $Y_1,...,Y_T$ دنباله غیرتصادفی با حد C می‌باشد، می‌توانیم مطمئن باشیم که $Y_1,...,Y_T$ که $|Y_t-c|<\varepsilon$ است وقتی T به اندازه کافی بزرگ می‌شود. وقتی $Y_1,...,Y_T$ دنباله تصادفی با حد C می‌باشد. فقط می‌گوییم $|Y_T-c|<\varepsilon$ با احتمال دلخواهی که نزدیک به یک است صحیح است. حد معمولی حالت خاصی از حد احتمال است.

مثال: دنباله تصادفی $x_1, x_2,..., x_n$ را در نظر بگیرید که دارای میانگین μ و واریانس σ^2 با توزیع نرمال است.

میانگین نمونه
$$\bar{X}_n=\frac{\sum X_i}{N}$$
$$Var(\bar{X})=\frac{\sigma^2}{N}$$

آنگاه $\bar{X}_n$ به‌طور احتمالی به سمت μ میل می‌کند.

$$Lim p\left[\left|\bar{X}-\mu\right|>\varepsilon\right]=0 \qquad \text{for} \quad \varepsilon>0$$

یا:
$$P\lim\bar{X}=\mu$$

اثبات از طریق قضیه نامساوی چه بی‌شف

این قضیه می‌گوید برای هر متغیر تصادفی Z با میانگین معین μ و واریانس σ^2، احتمال انحراف از میانگین برابر یا بزرگتر از K ضربدر انحراف معیارش حداکثر برابر $\frac{1}{K^2}$ است.

یعنی :
$$P\left[|Z-\mu|\geq K\sigma\right]\leq\frac{1}{K^2} \quad K>0$$

در مورد مثال $\bar{x}_n$ میانگین μ و واریانس $\frac{\sigma^2}{n}$ می‌باشد به‌طوریکه نامساوی فوق به‌صورت زیر خواهد شد.

$$\left|\bar{X}-\mu\right|\geq K\frac{\sigma}{\sqrt{n}}$$

یعنی:
$$K=\varepsilon\frac{\sqrt{n}}{\sigma}$$
$$\Rightarrow P\left[\left|\bar{X}_n-\mu\right|\geq\varepsilon\right]\leq\frac{\sigma^2}{n\varepsilon^2}$$

چون $\frac{\sigma^2}{n}$ در حد بسمت صفر میل می‌کند. در نتیجه:

$$Lim P\left[\left|\bar{X}_n-\mu\right|\geq\varepsilon\right]=0$$

نتیجه بسیار مهم

۳-۳ قضیه خین چین[13]

اگر $x_1, x_2, ..., x_n$ متغیری با توزیع مستقل و یکسان[14] و با میانگین معین μ باشد در آنصورت متوسط $x\left(\bar{x}_n\right)$، در احتمال به سمت μ میل می‌کند[15].

توجه کنیم این قضیه راجع به واریانس x سخنی به میان نمی‌آورد

۴-۳ قضایای اسلاتسکی

نتایج فوق قابل تعمیم به بردارها و ماتریسها می‌باشد همچنین اگر دنباله تصادفی $Y_1, ..., Y_n$ دارای حد احتمالی C می‌باشد، تابعی پیوسته از Y_i نیز به‌سمت تابعی ا ز C میل می‌کند.

$$\underset{n\to\infty}{\mathrm{Plim}}\, g(y_n) = g(\underset{n\to\infty}{\mathrm{Plim}}(y_n)) = g(c)$$

موارد استفاده:

$$\mathrm{Plim}(a_t + b_t) = \mathrm{Plim}\, a + \mathrm{Plim}\, b$$

$$\mathrm{Plim}\, a_t / b_t = \mathrm{Plim}\, a / \mathrm{Plim}\, b$$

$$\mathrm{Plim}\, a_t\ b_t = \mathrm{Plim}\, a . \mathrm{Plim}\, b$$

$$\mathrm{Plim}\, e^{a_t} = e^{\mathrm{Plim}\, a_t}$$

$$\mathrm{Plim}\, A_t^{-1} = (\mathrm{Plim}\, A_t)^{-1}$$

۵-۳ سازگاری[16]

حال چنانچه بجای دنباله تصادفی، تخمین زننده‌ای را در دست داشته باشیم مانند $\hat{\theta}$ که برآوردی از θ پارامتر جامعه می‌باشد اگر $\hat{\theta}_n$ در احتمال به‌سمت θ گرایش پیدا کند داریم:

$$\mathrm{Lim}\, p\left[\left|\hat{\theta}_n - \theta\right| > \varepsilon\right] = 0 \quad \text{for} \quad \varepsilon > 0$$

یا:

$$\underset{n\to\infty}{\mathrm{Plim}}\, \hat{\theta}_n = \theta$$

تخمین زننده فوق را سازگار می‌گویند.

مثلاً در مورد مثال میانگین نمونه گفتیم $\bar{x}$ در احتمال به‌سمت μ میل می‌کند، این جمله برابر است یا اینکه بگوییم $\bar{x}$ تخمین سازگاری از μ می‌باشد و این مسئله با نشان دادن اینکه $E(\bar{x}) = \mu$ و $\frac{\sigma^2}{n} \to 0$ روشن شد. (براساس قضیه چه بی‌شف)

پس: شرط کافی برای آن که $\hat{\theta}$ سازگار باشد آنست که در حد بدون تورش باشد و واریانس آن به سمت صفر میل کند:

$$\mathrm{Lim}\, E(\hat{\theta}) = \theta \qquad \text{و} \qquad \mathrm{Lim}\ \mathrm{Var}\,(\hat{\theta}) = 0$$

---

13 - Khintchine’s Theorem.
14 - Independently Identically Distributed.
۱۵- اثبات این قضیه از طریق تابع گشتاور میسر می‌باشد.
16 - Consistency.

توجه کنیم شرط کافی است ولی لازم نیست یعنی ممکن است تخمین زننده دارای تورش باشد ولی سازگار باشد یعنی تورش در حد به سمت صفر میل می‌کند.

در بعضی موارد حتی لازم نیست واریانس آن معین باشد: قضیه Khintchine می‌گوید اگر $x_1, x_2, ..., x_n$ متغیر i.i.d با میانگین معین μ باشد، درآنصورت $\overline{X}_n$ تخمین سازگاری از μ است.

۳-۵-۱ سازگاری $\hat{\beta}_{OLS}$ تحت شرایط استاندارد مدل رگرسیون خطی

در حقیقت سازگاری می‌گوید وقتی حجم نمونه بزرگ می‌شود، شانس بسیار کمی وجود دارد که تخمین یک پارامتر متفاوت از خود پارامتر باشد.

$$\hat{\beta}=\beta+(X'X)^{-1}X'U$$

با توجه به فرض اینکه:

$$\text{Lim}\frac{X'X}{T}=Q$$

داریم:

$$E(\hat{\beta})=\beta$$

$$\text{Var}(\hat{\beta})=\sigma^2(X'X)^{-1}$$

$$\text{Lim}\,\text{Var}(\hat{\beta})=\underbrace{\frac{\sigma^2}{T}}_{0}\underbrace{(\frac{X'X}{T})^{-1}}_{Q^{-1}}=0$$

لذا چون $\hat{\beta}$ بدون تورش است و واریانس آن به‌سمت صفر میل می‌کند تخمین زننده‌ای سازگار است.

۳-۵-۲ سازگاری $\hat{\sigma}^2$

$$\hat{\sigma}^2=\frac{\hat{U}'\hat{U}}{T-K}$$

$$E(\hat{\sigma}^2)=\sigma^2$$

$$\text{Var}(\hat{\sigma}^2)=\frac{2\sigma^4}{T-K}$$

چون واریانس آن به سمت صفر میل می‌کند و بدون تورش است، سازگار است[17].

۳-۵-۳ توزیع احتمالی[18] برای تخمین‌زننده‌های OLS

قبلاً در قضیه گوس و مارکف $\hat{\beta}$ را به‌صورت زیر معرفی کردیم:

$$\hat{\beta}=\beta+(X'X)^{-1}X'U$$

لذا اختلاف $\hat{\beta}-\beta$، تابعی خطی از u_t می‌باشد پس:

اگر از این عبارت امید بگیریم امید این اختلاف صفر است. یعنی $E(\hat{\beta})=\beta$

---

17 - $(T-K)\frac{\hat{\sigma}^2}{\sigma^2}\sim\chi^2_{T-K} \Rightarrow E(\chi^2_{T-K})=T-K \Rightarrow \text{Var}(\chi^2_{T-K})=2(T-K)$

18 - Probability Distribution.

و اگر:

$$u_i \sim N(0,\sigma^2)$$

در نتیجه:

$$\hat{\beta} \sim N(\beta,\sigma^2 (X'X)^{-1})$$

اما معمولاً σ^2 معلوم نیست و به‌جای آن از $\hat{\sigma}^2$ استفاده می‌شود. $\hat{\sigma}^2$ به‌صورت زیر است:

$$\hat{\sigma}^2 = \frac{\sum \hat{u}_t^2}{T-K} = \frac{\hat{U}'\hat{U}}{T-K}$$

قبلاً ثابت کردیم $\hat{u} = My = Mu$ پس

$$\hat{U}'\hat{U} = Y'MY = U'MU \qquad (۱)$$

عبارت (۱) فرم درجه دوم برحسب u_t می‌باشد که ماتریس وزنی آن M است که متقارن و هم قوه است.

به‌طوریکه:

$$Trace(M) = T-K = Rank(M)$$

چون M متقارن است لذا این امکان وجود دارد که ماتریس متعامد C را طوری پیدا کرد که $C'MC = D_{T-K}$ باشد (D ماتریس قطری است که T-K یک و K صفر درقطر اصلی دارد و C ماتریس بردارهای ویژه ماتریس M می باشد.)

بردار u را می توانیم بوسیله C به بردار V تبدیل کنیم.

$$V = C'U \Rightarrow U = CV$$
$$E(VV') = E(C'UU'C) = C'\sigma^2 IC$$
$$= \sigma^2 C'C$$

در $\hat{U}'\hat{U} = U'MU$ برحسب V جایگزین می کنیم:

$$\hat{U}'\hat{U} = U'MU = V'C'MCV$$
$$= V'D_{T-K}V$$
$$= V_1^2 + V_2^2 + \ldots + V_{T-K}^2 \qquad (۲)$$

الف- براساس عبارت (۲): اگر U_t ها بطور مستقل و نرمال توزیع شده‌اند بامیانگین صفر و واریانس σ^2، V_t ها هم به همین ترتیب توزیع شده اند.

ب- مجموع مجذورات پسماندها عبارت از مجموع مجذورات $T-K$، V_t که به طورمستقل و نرمال توزیع شده است، می باشد.

پس:

$$(T-K)\frac{\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{T-K} \qquad (۳)$$

برای آنکه نشان دهیم (۳) از $\hat{\beta}$ بطور مستقل توزیع شده است می باید نشان دهیم که $\hat{U}$ بطور مستقل از $\hat{\beta}$ توزیع شده است. چون $\hat{u}$ و $\hat{\beta}$ هر کدام تابعی خطی از توزیع نرمال هستند. مستقل بودن ازرابطه زیر بدست می‌آید:

$$E\left[\hat{u}(\hat{\beta}-\beta)'\right] = E\left[\left(I_T - X(X'X)^{-1}X'\right)uu'x(x'x)^{-1}\right]$$
$$= \sigma^2 X(X'X)^{-1} - \sigma^2 (X(X'X)^{-1} = 0$$

همانطور که قبلاً گفته شد هرگاه $T \rightarrow \infty$ میل کند،

$$Cov(\hat{\beta}) \rightarrow 0$$

تحت شرط $\mathrm{Lim}\dfrac{X'X}{T}=Q$، به این نتیجه می رسیم که وقتی $T\rightarrow\infty$ ، $\hat{\beta}$ توزیعش به نقطه تبدیل می‌شود که اصطلاحاً می گوئیم degenerate می شود.

لذا دربحث توزیع احتمالی ، $\hat{\beta}-\beta$، را در مقیاس $\sqrt{T}$، ضرب نموده تا درآنصورت $\sqrt{T}(\hat{\beta}-\beta)$ به سمت صفر میل نکند.

چون: $\mathrm{Var}(\sqrt{T}(\hat{\beta}-\beta))=\sigma^2T(X'X)^{-1}\neq 0$

پس راه دیگر بیان $\hat{\beta}\sim N(\beta,\sigma^2\,(X'X)^{-1})$ آن است که بگوئیم: $\sqrt{T}(\hat{\beta}-\beta)$ دارای توزیع چند بعدی با میانگین صفر و واریانس کوواریانس $\sigma^2(\dfrac{X'X}{T})^{-1}$ است.

پس اگر u_tها نرمال نباشند، توزیع دقیق $\hat{\beta}$ را مشکل بتوان بدست آورد. در چنین شرایطی راهی که موجب بدست آوردن توزیع احتمالی $\hat{\beta}$ بطور تقریب می‌شود آن است که ویژگی های حدی $\sqrt{T}(\hat{\beta}-\beta)_{T\rightarrow\infty}$ را در نظر گرفت.

تحت شرایط فروض کلاسیک $\sqrt{T}(\hat{\beta}-\beta)\rightarrow N(0,\sigma^2Q^{-1})$ این توزیع را توزیع مجانبی می‌گویند، $\sqrt{T}$را سرعت همگرایی می‌نامند و به تخمین زننده‌های OLS، $\sqrt{T}$سازگار اطلاق می‌شود، چون سرعت آن $\sqrt{T}$است

فصل چهارم

فصل چهارم

فصل چهارم

۴-۱ معنی‌دار بودن ضرایب و فاصله اعتماد

غالباً آزمونی که در مدل رگرسیون خطی برای ضرایب انجام می‌شود این است که آیا ضرایب برابر مقدار خاصی می‌باشند یا خیر؟ مثلاً این مقدار خاص می‌تواند صفر باشد.

β_j ضریب متغیر j ام مدل را در نظر بگیرید فرضیه صفر این است که: $H_\circ : \beta_j = 0$

آزمون کردن این فرضیه یعنی اینکه آزمون نمودن این موضوع که آیا متغیر j ام از نظر آماری بطور معنی‌داری روی متغیر وابسته اثر می‌گذارد یا خیر؟

تحت فرض نرمال بودن جمله اخلال ضرایب تخمین‌زده شده، دارای توزیع نرمال خواهد بود و بطور خاص $\hat{\beta}_j$ دارای توزیع نرمال با میانگین β_j و واریانس $\sigma^2 (X'X)^{-1}_{jj}$ خواهد بود، که $\sigma^2 (X'X)^{-1}_{jj}$ عبارتست از عنصر j ام روی قطر ماتریس VC .

$$1) \quad \frac{\hat{\beta}_j - \beta_j}{\sqrt{\sigma^2 (X'X)^{-1}_{jj}}} \sim Z$$

بطور استاندارد نرمال توزیع شده است.

اما می‌دانیم که تخمین σ^2 یعنی $\hat{\sigma}^2$، بصورت زیر توزیع شده است.

$$\hat{\sigma}^2 \sim \frac{\sigma^2}{T-K} \chi^2_{(T-K)}$$

لذا نسبت (۱) را بصورت نسبت (۲) می‌نویسیم که دارای توزیع t می‌باشد.

$$2) \quad \frac{(\hat{\beta}_j - \beta_j) \Big/ \sqrt{\sigma^2 (X'X)^{-1}_{jj}}}{\sqrt{\hat{\sigma}^2 / \sigma^2}} \sim t_{T-K}$$

چون j ام استاندارد ارور خواهد شد.

$$\hat{\sigma}_j=\sqrt{\hat{\sigma}^2\,(X'X)^{-1}_{jj}}\quad\sqrt{\frac{1}{(T-K)}\sum\hat{u}^2_t\;(X'X)^{-1}_{jj}}$$

لذا خواهیم داشت:

$$\frac{\hat{\beta}_j-\beta_j}{\hat{\sigma}_j}\sim t_{T-K}$$

براساس نظریه احتمال

$$P(-t_{\alpha/2}<t<t_{\alpha/2})=1-\alpha$$

بنابراین :

$$P\left[\hat{\beta}_j-t_{\alpha/2}.\hat{\sigma}_j<\beta_j<\hat{\beta}_j+t_{\alpha/2}.\hat{\sigma}_j\right]=1-\alpha$$

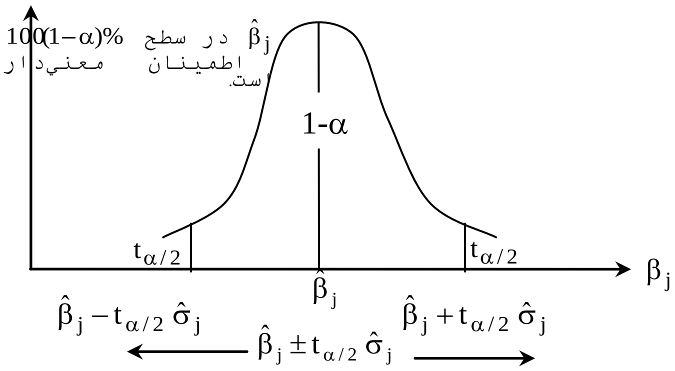
بطوریکه $t_{\alpha/2}$ مقدار توزیع t برای سطح معنی دار بودن $\alpha/2$ با درجه آزادی T-K است. پس با احتمال $100(1-\alpha)\%$ فاطمه اعتماد برای ضریب $\hat{\beta}_j\pm-t_{\alpha/2}.\hat{\sigma}_j$ خواهد شد

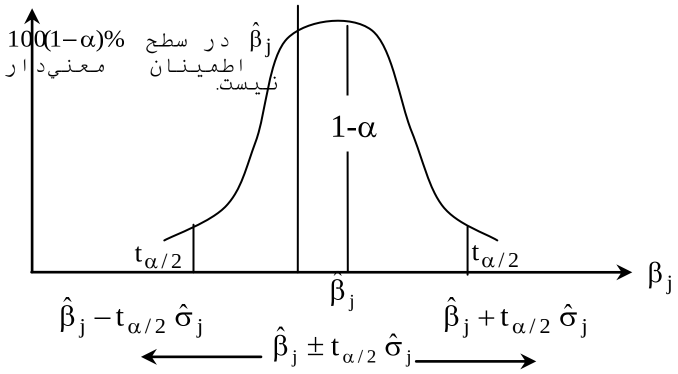
مثلاً ۹۵٪ فاصله اعتماد عبارت است از تخمین پارامتر j ام و $\hat{\beta}_j$ به اضافه یا منهای مقدار توزیع t در سطح معنی‌دار بودن 025. ضربدر انحراف معیار مربوطه.

ضریب $\hat{\beta}_j$، بطور معنی‌داری از صفر متفاوت است هرگاه فاصله اعتماد شامل صفر نشود. (فرض $H_\circ$ را رد می‌کنیم) بعکس پارامتر $\hat{\beta}_j$ بطور معنی‌داری از صفر متفاوت نیست، هرگاه فاصله اعتماد مقدار صفر را شامل شود. ($H_\circ$ را می‌پذیریم)

فرض کنیم $\hat{\beta}_j$مثبت است، اگر فاصله اعتماد مرکز مختصات را در بر بگیرد در آنصورت نسبت t برای ضریبjام بصورت $t_j=\dfrac{\hat{\beta}_j}{\sigma_j}$ خواهد شد.

در شکل توجه شود.





لذا نسبت t_j عبارتست از نسبت تخمین ضریب رگرسیون به انحراف معیارش، که معنی دار بودن ضریب را تعیین می‌کند: بطور کلی فرضیه صفر $H_\circ : \beta_j = 0$ پذیرفته می‌شود اگر قدر مطلق $|t_j|$ کوچکتر ازمقدار t متناظر با سطح خاصی از معنی‌دار بودن باشد و این فرضیه رد می‌شود در صورتیکه $|t_j|$ از این مقدار بیشتر باشد.

نسبت t کوچک نشان می‌دهد که متغیر توضیحی مربوط رابطه‌ای با متغیر وابسته مدل ندارد. ولی اگر نسبت t از مقدار بحرانی(درسطح انتخاب شده معنی دار بودن) بزرگتر باشد، در آنصورت ضریب مربوط معنی دار است، یعنی متغیر وابسته بطور خطی بامتغیر توضیحی مربوط وابسته است.

برای درجه آزادی بزرگ ($T > 30$) ، توزیع t تقریباً شبیه نرمال است در این مورد قاعده کلی[۱۹] برای تفسیر این است که اگر نسبت t بزرگتر از2 می‌باشد، ضریب مربوط معنی دار است و بر عکس، اگر نسبت t از۲ کوچکتر است درآنصورت ضریب مربوط بطور آماری با صفر تفاوتی ندارد. نسبت t، می‌تواند برای آزمون مقادیر دیگر از $\hat{\beta}_j$ و سایر آزمونهای مثل F و غیره مورد استفاده قرار گیرد.

حال فرضیه صفر زیر را در نظر بگیرید: $H_\circ : \beta_j = \beta_j^\circ$

فرض کنید $\hat{\beta}_j$ از β_j° بزرگتر است، ممکن است فاصله اعتماد طوری انتخاب شود که β_j° را بپوشاند.

۱۹- مثل قانون شصت در تعیین جهت میدان مغناطیسی در فیزیک

$$\hat{\beta}_j - t_{\alpha/2}.\hat{\sigma}_j = \beta_j$$

و t خواهد شد:

$$t = \frac{\hat{\beta}_j - \beta_j^{\circ}}{\hat{\sigma}_j}$$

بطور کلی، اگر قدر مطلق t از $t_{\alpha/2}$ برای درجه آزادی T-K ، بزرگتر باشد، آنگاه سطح معنی‌دار بودن $100(1-\alpha)\%$ فرضیه رد می‌شود، اگر |t| پایین منطقه بحرانی بیفتد فرضیه $H_{\circ}$ پذیرفته می‌شود.

فصل پنجم

فصل پنجم

۱-۵ پیش‌بینی[۲۰]

همانطور که گفته شد وظیفه اصلی اقتصاد سنجی پیش‌بینی است، که بر اساس تخمین پارامترهای مدل رگرسیون انجام میگیرد. فرض کنیم محققی علاقه مند است، مقدارY_{T+1} را متناظر با بردار رگرسور X_{T+1} پیش‌بینی کند. مقدار Y در دوره $T+1$ عبارت است از:

$$Y_{T+1} = X_{T+1}\beta + u_{T+1}$$

با توجه به قضیه گوس- مارکف و با جایگذاری $\hat{\beta}$ که تخمین زننده ای خطی و بدون تورش است $y^p{}_{T+1} = X_{T+1}\hat{\beta}$ خواهد شد که پیش بینی کننده خطی بدون تورش با کمترین واریانس برای y_{T+1} است چون:

$$E(y^p{}_{T+1}) = X_{T+1}\beta$$

اما در اینجا دچار خطای پیش‌بینی خواهیم شد که عبارتست از تفاضل $y_{T+1} - y_{T+1}{}^p$ که y_{T+1} مقدار تحقق یافته y در دوره پیش‌بینی شده است.

$$e^p = y_{T+1} - y_{T+1}{}^p = X_{T+1}\beta + u_{T+1} - X_{T+1}\hat{\beta}$$
$$e^p = X_{T+1}(\beta - \hat{\beta}) + u_{T+1}$$

در اینجا خطای ما بطور متوسط صفر است چون:

$$E(e^p) = o$$

واریانس خطای پیش‌بینی کننده عبارتست از:

---

20. Prediction

$$Var(e^p) = \sigma^2 + Var[X_{T+1}(\beta - \hat{\beta})]$$
$$= \sigma^2 + X_{T+1}[\sigma^2 (XX')^{-1}]X'_{T+1}$$

اگر مدل دارای عرض از مبدا باشد در آنصورت:

$$Var(e^p) = \sigma^2[1 + \frac{1}{T} + \sum_{j=1}^{k-1}\sum_{k=1}^{k-1} (x_{T+1,j} - \bar{x}_j)(x_{T+1,k} - \bar{x}_k)(z'M^0 z)^{jk}]$$
$$Var(e^p) = \sigma^2\{1 + \frac{1}{T} + [(x_1 - \bar{x}_1)(z'M^0 z)^{11} + (x_2 - \bar{x}_2)(z'M^0 z)^{22} + \ldots + (x_{k-1} - \bar{x}_{k-1})(z'M^0 z)^{k-1}]^2\}$$

بطوریکه Z عبارتست از k-1 ستون x ، یعنی:

$$\underset{T.k}{X} = [\underset{T.1}{1} \ \underset{T.k-1}{z}]$$

و $M^0 = I - i(i'i)^{-1}i'$ ، بطوریکه: $i' = [1,1,\ldots,1]'_{1.T}$

این نتیجه نشان میدهد که میزان پراکندگی و خطای پیش بینی کننده بستگی به فاصله x_{T+1} از مرکز مشاهدات دارد و با حجم مشاهدات بطور معکوس رابطه دارد، یعنی درجه نااطمینانی در پیش‌بینی با کاهش اندازه نمونه افزایش می‌یابد.

$Var(e^p)$ با استفاده از $\hat{\sigma}^2$ بجای σ^2 قابل محاسبه است.

فاصله اعتماد نیز برای y^p_{T+1} بصورت زیر خواهد بود:

فاصله اعتماد برای پیش‌بینی کننده = $y^P_{T+1} \pm t_{\alpha/2}.Se(e^p)$

$$Se(e^p) = \sqrt{Var(e^p)}$$

مثال:

پیش‌بینی میزان سرمایه‌گذاری حقیقی برای مدل سرمایه‌گذاری که قبلا به آن پرداخته شد برای یکسال بعد بصورت زیر است:

X_{T+1} = (1, 16, 1.5, 10, 4)
1 =Intercept
16 = T + 1 = 1983
1.5 = Real GNP
10 = Interest Rate
4 = Inflation Rate

با استفاده از نتایج مدل رگرسیون زیر که برای t = 1962 - 1982 محاسبه شده است:

Real Invt = $\beta_0 + \beta_1$T + β_2 Real GNP + β_3 Interest Rate + β_4 Inflation Rate

Real Invt = -0.509 – 0.017T + 0.67 RGNP – 0.002 IntR – 0.00009 InfR

$$(\mathrm{Re}\,\hat{alInv})^{p} = (1,16,1.5,10,4)[-0.509,-0.017,0.67,-0.002,0.0001]'$$

=0.2036

تخمین واریانس این پیش‌بینی کننده عبارتست از:

$$\hat{\sigma}^2[1+X_{T+1}(X'X)^{-1}X_{T+1}'] = 0.000097$$

و با جذر گرفتن از عدد فوق فاصله اطمینان برای پیش‌بینی کننده عبارت خواهد شد از:

$$0.2036 \pm 2.228(0.009885) = (0.1811 \leftrightarrow 0.2262)$$

فصل ششم

فصل ششم

۶-۱ روش حداقل مربعات همراه با محدودیت[21]

اجازه دهید موضوع را با چند مثال اقتصادی که با آن آشنا هستید آغاز کنیم.

مثال۱: تابع تولید کاب-داگلاس زیرا در نظر بگیریم. میدانیم که همگن درجه اول بودن شرط مناسبی برای تولید است. اما سئوال این است که این شرط در داخل داده‌های مورد مشاهده یک بنگاه وجود دارد؟ برای پاسخ به چنین سوالی ابتدا با لگاریتم گیری تابع را به صورت خطی تبدیل میکنیم. در این تابع Q سطح تولید K نهاده سرمایه و L نهاده کار است. A ضریب تکنولوژی و u جمله اختلال میباشد.

$$Q_t = AK_t^{\alpha} L_t^{\beta} e^{u_t}$$
$$LnQ_t = LnA + \alpha Lnk_t + \beta LnL_t + u_t$$
$$Y_t = \beta_1 + \beta_2 X_{t1} + \beta_3 X_{t2} + u_t$$
$$\beta_1 = LnA \qquad X_{t1} = LnK_t$$
$$\beta_2 = a \qquad X_{t2} = LnL_t$$
$$\beta_3 = \beta$$

فرض کنیم $\beta_2 + \beta_3 = 1$ یعنی همگن از درجه اول بودن تابع تولید. با اعمال چنین قیدی در تابع مزبور داریم:

$$Y_t = \beta_1 + \beta_2 X_{t1} + (1 - \beta_2) X_{t2} + u_t$$
$$Y_t - X_{t2} = \beta_1 + \beta_2 (X_{t1} - X_{t2}) + u_t$$
$$Y^* = \beta_1 + \beta_2 X^* + u_t$$

21. Restricted OLS

ملاحظه میشود که مدل فوق تابعی است صرفا با دو پارامتر.

مثال ۲: مدل ۵ متغیره زیررا در نظر بگیریم.

$$Y = X\beta + \varepsilon$$

بطوریکه :

$$\beta' = (\beta_1, \beta_2, \beta_3, \beta_4, \beta_5)$$

چنانچه بین پارامترهای مدل بالا روابط زیر برقرار باشد:

$$\beta_1 - 3\beta_2 + \beta_5 = 15$$
$$\beta_1 + \beta_3 = 0$$

روابط فوق را که بر اساس نظریه‌های اقتصادی لحاظ میشود را میتوانیم در قالب فرم ماتریسی بنویسیم.

ماتریس R و بردار r را در نظر بگیریم :

$$R = \begin{bmatrix} 1 & -3 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \end{bmatrix}$$

$$\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_5 \end{bmatrix} \quad r = \begin{bmatrix} 15 \\ 0 \end{bmatrix}$$

قید ما در این مدل عبارتست از:

$$R\beta = r$$

R ماتریسیG.K و β بردار K.1 وr برداری1. G میباشد. حال در بکار بردن روش حداقل مربعات میباید قید بالا لحاظ شود.

لـــــــذا:

$$Min \sum \hat{u}_t^2 = \hat{u}'\hat{u}$$
$$S.t\ R\beta^* = r$$

و یا :

$$\begin{aligned}
&\mathrm{Min}\,(Y-X\beta^*)'(Y-X\beta^*)\\
&\quad \mathrm{S.t}\, R\beta^* = r\\
&L=(Y-X\beta^*)'(Y-X\beta^*)+2\lambda'\left[r-R\beta^*\right]\\
&\;=Y'Y-2Y'X\beta^*+\beta^{*\prime}X'X\beta^*+2\lambda'\left[r-R\beta^*\right]\\
&\frac{\partial L}{\partial \beta^*}=-2X'Y+2X'X\beta^*+2R'\lambda^*=0 \qquad (1)\\
&\frac{\partial L}{\partial \lambda}=-2r+2R\beta^*=0 \qquad (2)\\
&\quad\Rightarrow\quad r=R\beta^*\\
&\quad\Rightarrow\quad X'X\beta^*-X'Y+R'\lambda^*=0\\
&(X'X)^{-1}X'X\beta^*=(X'X)^{-1}X'Y+(X'X)^{-1}R'\lambda^*\\
&\beta^*=(X'X)^{-1}X'Y+(X'X)^{-1}R'\lambda^*
\end{aligned}$$

اگر طرفین رابطه بالا را در R پیش ضرب می‌کنیم داریم:

$$R\beta^*=R\hat{\beta}+R(X'X)^{-1}R'\lambda^*$$

حال چنانچه محدودیت را برای مقادیر تخمین زده شده اعمال کنیم $R\beta^*=r$ خواهد شد لــــذا:

$$r=R\hat{\beta}+\left[R(X'X)^{-1}R'\right]\lambda^*$$

$$r-R\hat{\beta}=\left[R(X'X)^{-1}R'\right]\lambda^*$$

معادله فوق را برای λ^* حل کنیم:

$$\lambda^*=\left[R(X'X)^{-1}R'\right]^{-1}(r-R\hat{\beta})$$

و در نتیجه خواهیم داشت:

$$\beta^*=\hat{\beta}+(X'X)^{-1}R'\left[R'(X'X)^{-1}R'\right]^{-1}(r-R\hat{\beta})$$

چنانچه محدودیتها در داخل نمونه وجود دارند آنگاه $R\hat{\beta}=R\beta=r$ خواهد شد و در نتیجه: $\beta^*=\hat{\beta}$

بدون تورش بودن $\overset{*}{\beta}$: می توان بسادگی نشان داد که β^* بدون اریب است.

$$\begin{aligned}
&\hat{\beta}=\beta+(X'X)^{-1}X'\varepsilon\\
&\beta^*=\beta+(X'X)^{-1}X'\varepsilon+(X'X)^{-1}R'\left[R(X'X)^{-1}R'\right]^{-1}\left[\left(r-R(\beta+(X'X)^{-1}X'\varepsilon\right)\right]\\
&\beta^*=\beta+(X'X)^{-1}X'\varepsilon-(X'X)^{-1}R'\left[R(X'X)^{-1}R'\right]^{-1}\left[R(X'X)^{-1}X'\varepsilon\right]\\
&=\beta+\left[I-(X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}R\right](X'X)^{-1}X'\varepsilon\\
&E(\beta^*)=\beta
\end{aligned}$$

ماتریس واریانس، کوواریانس β^*

$$\begin{aligned}
Vc\ (\beta^*) &= E(\beta^* - \beta)(\beta^* - \beta)' \\
&= E\{I - X'X)^{-1} R'[R(X'X)^{-1} R']^{-1} R\}(X'X)^{-1} X'\varepsilon \\
&\quad \varepsilon' X(X'X)^{-1}\{I - (X'X)^{-1} R'[R(X'X)^{-1} R']R\}' \\
&= \sigma^2 (X'X)^{-1}\{I - R'[R(X'X)^{-1} R']^{-1} R(X'X)^{-1}\} \\
&= Vc(\hat{\beta}) - Vc(\hat{\beta}) R'[R\ Vc(\hat{\beta}) R']^{-1} R\ \ Vc(\hat{\beta})
\end{aligned}$$

چون واریانس $\hat{\beta}$ مثبت معین است و R دارای مرتبه کامل است، لذا $R\,Vc(\hat{\beta})\,R'$ مثبت معین است در نتیجه $[R\,Vc(\hat{\beta})\,R']$ مثبت معین و معکوس پذیر است[22].

در نتیجه:

$$Vc(\hat{\beta})\,R'[R\,Vc(\hat{\beta})\,R']^{-1} R\,Vc(\hat{\beta})$$

مثبت شبه معین است. لذا تفاوت بین $Vc(\hat{\beta})$ و $Vc(\beta^*)$ مثبت شبه معین است، و این نشان میدهد که عناصر روی قطر ماتریس واریانس – کوواریانس β^* کمتر یا مساوی عناصر متناظر از ماتریس واریانس – کوواریانس $\hat{\beta}$ می باشد. این نتایج همواره برقرار است.

22.

$$H_0 : R\beta = r$$

$$\overset{*}{\beta} = \hat{\beta} - (X'X)^{-1}[R(X'X)^{-1}R]^{-1}(R\hat{\beta} - r)$$

$$E(\overset{*}{\beta}) = \beta$$

$$\begin{aligned}
VC(\overset{*}{\beta}) &= E\{I - (X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}R\}(X'X)^{-1}X'\varepsilon. \\
&\qquad \varepsilon' X(X'X)^{-1}\{I - (X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}R\}' = \\
&= \{I - (X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}\}(X'X)^{-1}X'\sigma^2 X(X'X)^{-1}\{I - (X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}R\}' = \\
&= \sigma^2\{I - (X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}R\}(X'X)^{-1}\{I - (X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}R\}' \\
&= \sigma^2\{(X'X)^{-1} - (X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}R(X'X)^{-1}\}\{I - (X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}R\}' \\
&= \sigma^2(X'X)^{-1}\{I - R'[R(X'X)^{-1}R']^{-1}R(X'X)^{-1}\}\{I - (X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}R\}' \\
&= \sigma^2(X'X)^{-1}\{I - R'[R(X'X)^{-1}R']^{-1}R(X'X)^{-1} - R'[R(X'X)^{-1}R']^{-1}R(X'X)^{-1} + \\
&\qquad R'[R(X'X)^{-1}R']^{-1}[R(X'X)^{-1}R'][R(X'X)^{-1}R']^{-1}R(X'X)^{-1}\} \\
&= VC(\hat{\beta})\{I - 2R'[R(X'X)^{-1}R']^{-1}R(X'X)^{-1} + R'[R(X'X)^{-1}R']^{-1}R(X'X)^{-1}\} \\
&= VC(\hat{\beta})\{I - R'[R(X'X)^{-1}R']^{-1}R(X'X)^{-1}\} \\
&= VC(\hat{\beta})\{I - R'[R\,VC(\hat{\beta})\,R']^{-1}R\,VC(\hat{\beta})\}
\end{aligned}$$

$$VC(\overset{*}{\beta}) = VC(\hat{\beta}) - \{VC(\hat{\beta})R'[R\,VC(\hat{\beta})\,R']^{-1}\}R\,VC(\hat{\beta})\}$$

قضیه: مجموع مجذورات پسماندهای مدل مقید بزرگتر از مجموع مجذورات مدل بدون قید است.

$RSS_R \geq RSS_u$ یا $\hat{u}^{*\prime}\hat{u}^{*} \geq \hat{u}'\hat{u}$

$\hat{u}^{*\prime}\hat{u}^{*} = (y - X\beta^{*})'(y - X\beta^{*})$

و می دانیم:

$$\beta^{*} = \hat{\beta} + (X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}(r - R\hat{\beta})$$
$$\hat{u}^{*\prime}\hat{u}^{*} = (y - XB^{*})'(y - XB^{*})$$
$$\beta^{*} = \hat{\beta} + (X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}[r - R\hat{\beta})$$
$$\hat{u}^{*\prime}\hat{u}^{*} = \{y - X\hat{\beta} - X(X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}[r - R\hat{\beta}]\}'*$$
$$= \{y - X\hat{\beta} - X(X'X)^{-1}R[R(X'X)^{-1}R']^{-1}[r - R\hat{\beta}]\}$$
$$\hat{u}^{*\prime}\hat{u}^{*} = y'y - y'x\hat{\beta} - y'x(X'X)^{-1}R']^{-1}[r - R\hat{\beta}] - \hat{\beta}x'y + \hat{\beta}x'x\hat{\beta}$$
$$+\hat{\beta}x'x(X'X)^{-1}R'[R(X'X)^{-1}R'](r - R\hat{\beta}R')][R'(X'X)^{-1}R]^{-1}R(X'X)^{-1}X'Y$$
$$+(r' - \hat{\beta}R')[R'(X'X)^{-1}R]^{-1}R(X'X)^{-1}X'X(X'X)^{-1}R'[R(X'X)^{-1}R]^{-1}R(X'X)^{-1}X'Y$$
$$+(r' - \hat{\beta}R')[R'(X'X)^{-1}R]^{-1}R(X'X)^{-1}X'X\hat{\beta}$$
$$\hat{u}^{*\prime}\hat{u}^{*} == y'y - y'x\hat{\beta} - \hat{\beta}X'X\hat{\beta} + (r' - \hat{\beta}R')[R'(X'X)^{-1}R]^{-1}(r - R\hat{\beta})$$
$$= (y - x\hat{\beta})'(y - x\hat{\beta}) + (r - R\hat{\beta})'[R'(X'X)^{-1}R]^{-1}(r - \hat{\beta})$$
$$\hat{u}^{*\prime}\hat{u}^{*} == \hat{u}'\hat{u} + (R\hat{\beta} - r)'[R(X'X)^{-1}R']^{-1}(R\hat{\beta})$$
$$\Rightarrow \hat{u}^{*\prime}\hat{u}^{*} => \hat{u}'\hat{u}$$
$$\Rightarrow RSS_R \geq RSS_u$$

۶-۲ استفاده از آزمون F برای بررسی قیدهای مدل رگرسیون خطـــی

هم چنانکه گفته شد مدل $Y = X\beta + U$ را با وجود قید $R\beta = r$ در نظر می‌گیریم، فرضیه صفر عبارتست از:

$H_0: R\beta = r$

ابتدا $R\hat{\beta}$ را تشکیل می‌دهیم. هر چقدر این بردار از r فاصله بگیرد، شک ما را درباره این فرضیه بیشتر می‌کند. لذا نیاز به توزیع نمونه ای $R\hat{\beta}$ داریم:

$E(R\hat{\beta}) = R\beta$

$Vc(R\hat{\beta}) = E[R(\hat{\beta} - \beta)(\hat{\beta} - \beta)'R'] = \sigma^2 R(X'X)^{-1}R'$

با فرض توزیع نرمال برای جمله اخلال مدل، $\hat{\beta}$ دارای توزیع چندگانه نرمال است.

لـــــذا:

$$R\hat{\beta} \sim N(R\beta, \sigma^2 R(X'X)^{-1}R')$$

یـــــا:

$$R(\hat{\beta}-\beta) \sim N(0, \sigma^2 R(X'X)^{-1}R')$$

اگر فرضیه $R\beta = r$ صحیح باشد می توانیم $R\beta$ را در رابطه فوق با r جایگزین کنیم.

پس:

$$R\hat{\beta}-r \sim N(0, \sigma^2 R(X'X)^{-1}R')$$

می‌دانیم چنانچه بردار تصادفی $Z_{n.1}$ دارای توزیع نرمال با میانگین صفر و ماتریس واریانس کوواریانس Σباشد. $Z'\Sigma^{-1}Z$ دارای توزیع χ^2 با درجه آزادی n میباشد.[۲۳]

در نتیجه :

$$(R\hat{\beta}-r)'[\sigma^2 R(X'X)^{-1}R']^{-1}(R\hat{\beta}-r) \sim \chi^2_{(G)}$$

G درجه آزادی است و تعداد سطرهای $R\hat{\beta}$ می‌باشد.

مشکل در اینجاست که σ^2 معلوم نیست، ولی قبلا نشان داده شده است[۲۴] که:

$$\frac{\hat{U}'\hat{U}}{\sigma^2} \sim \chi^2_{T-k}$$

و مستقل از $\hat{\beta}$ توزیع شده است و در نتیجه مستقل از $R\hat{\beta}$هم می‌باشد.

---

۲۳. اثبات : چونZ ها از یکدیگر مستقل نیستند، لذا می باید آنها را تبدیل به متغیر جدیدی نمود. چون Σ مثبت معین است لذا $\Sigma = p'p$ بطوریکه P ماتریس غیر منفرد است.

$$\Sigma^{-1} = (P^{-1})'P^{-1} \Rightarrow \quad P^{-1}\Sigma(P^{-1})' = I$$

$P^{-1}Z = Y$ را تعریف می کنیم بطوریکه:

$$E(Y) = P^{-1}E(Z) = 0$$

$$Var(Y) = E[P^{-1}ZZ'P^{-1}]P^{-1}\Sigma P^{-1'} = I$$

Y ها دارای توزیع استاندارد نرمال میباشند و داریم:

$$Y'Y \sim \chi^2_{(n)}$$

$$Y'Y \sim Z'(P^{-1})'P^{-1}Z = Z'\Sigma^{-1}Z \sim \chi^2_{(n)} \qquad \text{Q.E.D}$$

.۲۴

$$U \sim N(0, \sigma^2 I)$$

$$\hat{\beta} \sim N(\beta, \sigma^2 (X'X)^{-1}$$

$$\hat{U}'\hat{U} = U'MU; M = I - X(X'X)^{-1}X'$$

$$U'MU \sim \chi^2_{T-k}$$

T-K درجه آزادی است و برابر با رتبه و اثر ماتریس M می باشد.

لذا می توانیم توزیع F را بصورت زیر بسازیم، و σ^2نیز حذف می‌شود:

$$\frac{(R\hat{\beta}-r)'[R(X'X)^{-1}R']^{-1}(R\hat{\beta}-r)/G}{\hat{U}'\hat{U}/T-K}$$

اگر F محاسبه شده بیشتر از F جدول باشد (در سطح معین از احتمال) در آنصورت H_0 رد می‌شود. چنانچه F جدول کوچکتر باشد وجود قید را در داخل نمونه خواهیم پذیرفت. ملاحظه می‌شود که صورت توزیع F عبارت است از تفاضل مجموع مجذورات جملات خطا در رگرسیون مقید و غیرمقید، لذا می توانیم آنرا بصورت زیر بنویسیم:

$$\frac{(RSS_R-RSS_u)/G}{RSS_u/T-K}$$

اما چنانچه صورت و مخرج را بر TSS تقسیم کنیم داریم:

$$\frac{[(1-R_R^2)-(1-R_U^2]/G}{(1-R_U^2)/T-K}$$

$$\frac{(R_U^2-R_R^2)/G}{(1-R_U^2)/T-K}$$

لذا نحوه محاسبه F به مقتضای اطلاعات و داده های مدل تغییر می‌نماید.

۶-۳ آزمون وجود رابطه بین متغیرهای توضیحی و وابسته

چنانچه مدل رگرسیون خطی دارای عرض از مبدا باشد معمولا آزمون وجود رابطه بین y و متغیرهای مستقل انجام می گردد. بطوریکه:

$$H_0:\beta_2=\beta_3=\cdots=\beta_k=0 \qquad ; \qquad \beta_1=\text{عرض از مبدا}$$

$$y_t=\beta_1+u_t \qquad \text{تحت شرایط } H_0$$

$$\bar{y}=\beta_1^*$$

$$u_t^*=y_t-\bar{y}$$

$$RSS_R=\sum u_t^{2*}=\sum(y_t-\bar{y})^2=TSS_{yy}$$

$$RSS_u=\sum \hat{u}_t^2=TSS_{\hat{u}\hat{u}}$$

$$\frac{(RSS_R-RSS_u)/K-1}{RSS_u/T-K}=\frac{(TSS_{yy}-TSS_{\hat{u}\hat{u}})/K-1}{TSS_{\hat{u}\hat{u}}}$$

$$=\frac{S_{yy}/K-1}{(S_{yy}-S_{\hat{y}\hat{y}})/T-K}$$

صورت و مخرج را بر S_{yy} تقسیم می‌کنیم.

$$=\frac{(S_{yy}/S_{yy})/K-1}{(1-\frac{S_{\hat{y}\hat{y}}}{S_{yy}})/T-K}$$

داریم:

$$=\frac{R^2}{1-R^2}\cdot\frac{T-K}{K-1}\sim F(_{k-1},_{T-k})$$

F مزبور را که معمولا درگزارشهای رایانه‌ای مشاهده می‌کنیم ، نشانگر وجود رابطه بطور کلی بین متغیرهای توضیحی و متغیر وابسته می‌دانند.

۶-۴ آزمون برابری دو مدل رگرسیون در دو وضعیت خاص اقتصادی مانند جنگ و صلح

فرض کنیم T_1 مربوط به دوران جنگ و T_2 مربوط به دوران صلح باشد آنگاه دو مدل زیر را در نظر بگیریم:

$$y_1 = X_1\beta_1 + u_1 \qquad \text{جنگ} \qquad t\in T_1, \beta_1 : KX_1$$

$$y_2 = X_2\beta_2 + u_2 \qquad \text{صلح} \qquad t\in T_2, \beta_2 : KX_2$$

$$H_0 : \beta_1 = \beta_2 = \beta$$

تحت شرایط H_0 (مدل مقید) داریم:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}\beta + \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

از انجام رگرسیون فوق RSS_R بدست می‌آید.

تحت شرایط فرضیه مقابل (مدل غیر مقید) داریم:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix}\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

از انجام رگرسیون مزبور RSS_U بدست می آید.

$$T = T_1 + T_2$$

$$F = \frac{[RSS_R - (RSS_1 + RSS_2)]/K}{[RSS_1 + RSS_2]/T-2K} \sim F_{k}^{a}, {}_{T-2k}$$

چنانچه F محاسبه شده فوق بزرگتر از F جدول در سطح احتمال معین باشد فرضیه H_0 رد می‌شود. یعنی یکسان بودن ضرایب مدل در دو وضعیت اقتصادی رد می‌شود و چنانچه، کوچکتر از F جدول در سطح احتمال معین باشد، یکسان بودن ضرایب پذیرفته می‌شود.

مثال: حال با اطلاعات داده شده میخواهیم شرط همگن بودن از درجه اول را در تابع $Q=AK^{a}L^{\beta}e^{u}$ را که در ابتدای بحث اشاره شد را آزمون کنیم. میخواهیم فرضیه $H_0 = a+\beta=1$. ابتدا قید را در مدل اعمال نموده و آنرا بصورت مقید تخمین می زنیم. تعداد مشاهدات T =۳۰ می باشد.

$$\alpha = 1-\beta \qquad G=1 \qquad K=1$$

$$\begin{aligned} Ln\,Q &= LnA + (1-\beta)\,LnK + \beta\,LnL + u \\ &= LnA + LnK + \beta\,(LnL - LnK) + u \end{aligned}$$

$$Ln\,(Q/K) = LnA + \beta\,Ln\,(L/K) + u \qquad RSS_R = 1.2$$

$$Ln\,Q = LnA + \alpha\,LnK + \beta\,LnL + u \qquad RSS_u = 1$$

$$F = \frac{(RSS_R - RSS_{u})/1}{RSS_u/27} = \frac{(1.2-1)/1}{1/27} = 5.4$$

با مراجعه به جدول در سطح ۵٪ = α و درجه آزادی ۲۷ ملاحظه می‌کنیم ۴/۲۵ = F است لذا چون ۴/۲۵> ۵/۴ فرضیه همگن بودن از درجه اول برای تابع فوق رد می شود.

ولی چنانچه $\alpha = 1\%$ باشد، ۷/۸۲ = F و لذا بازدهی ثابت نسبت به مقیاس رد نمی شود.

۶-۵ آزمون تغییرات ساختاری Chow test (1960)

در تصریح یک مدل، فرض می کنیم که فروض ما در تمام نمونه صادق است. ولی منطقی به نظر می‌رسد که فرض کنیم در زیر مجموعه‌ای از مشاهدات ما فروض تغییر نماید و یا دارای ضرایب متفاوتی باشیم.

مثلا در بازار بنزین برای کشورهای غربی مجموعه داده‌های آماری نشان میدهد که رفتار بازار قبل و بعد از شوک نفتی 1973 تفاوت قابل ملاحظه‌ای نموده است. جهش قیمت درسال ۱۹۷۳ و ۱۹۸۰ کاملا هویداست و همچنین تغییرات زیادی در مصرف مشاهده میشود.

مدل زیر را در نظر بگیریم:

$$Log\,(G/POP)_t = \beta_1 + \beta_2 \log y_t + \beta_3 \log PG_t + \beta_4 \log PNC_t + \beta_5 \log PUC_t + \beta_6 t + \varepsilon_t$$

G: میزان بنزین مصرفی

POP: جمعیت

Y: درآمد سرانه

PG : قیمت G

PNC : قیمت اتومبیل جدید

PUC : قیمت اتومبیل دست دوم

با استفاده از داده‌های آماری کل نمونه ۱۹۹۵- ۱۹۶۰ و برای دو دوره ۱۹۷۳-۱۹۶۰ و ۱۹۹۵- ۱۹۷۴ سه رگرسیون انجام داده‌ایم که نتایج آن در جداول شماره ۱ -۲ -۳ آمده است.

جدول۱

Dependent Variable: LOG(G/POP)
Method: Least Squares
Date: 11/03/04 Time: 08:53
Sample: 1960 1995
Included observations: 36

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	20.06662	10.11464	1.983918	0.0565
LOG(Y)	1.954626	0.192854	10.13525	0.0000
LOG(PG)	-0.115530	0.033479	-3.450760	0.0017
LOG(PNC)	0.205282	0.152019	1.350368	0.1870
LOG(PUC)	-0.129274	0.071412	-1.810263	0.0803
YEAR	-0.019118	0.005957	-3.209614	0.0032

R-squared	0.968725	Mean dependent var	-0.003709
Adjusted R-squared	0.963512	S.D. dependent var	0.151691
S.E. of regression	0.028976	Akaike info criterion	-4.093705
Sum squared resid	0.025188	Schwarz criterion	-3.829785
Log likelihood	79.68669	F-statistic	185.8440
Durbin-Watson stat	0.778643	Prob(F-statistic)	0.000000

جدول ۲

Dependent Variable: LOG(G/POP)
Method: Least Squares
Date: 11/03/04 Time: 18:34
Sample: 1960 1973
Included observations: 14

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-55.78632	18.04991	-3.090670	0.0149
LOG(Y)	0.423995	0.351778	1.205291	0.2625
LOG(PG)	0.094547	0.150511	0.628172	0.5474
LOG(PNC)	0.583896	0.131493	4.440510	0.0022
LOG(PUC)	-0.334619	0.092326	-3.624299	0.0067
YEAR	0.026366	0.010690	2.466382	0.0389

R-squared	0.998033	Mean dependent var	-0.138297
Adjusted R-squared	0.996804	S.D. dependent var	0.159715
S.E. of regression	0.009030	Akaike info criterion	-6.279088
Sum squared resid	0.000652	Schwarz criterion	-6.005206
Log likelihood	49.95362	F-statistic	811.8469
Durbin-Watson stat	2.391035	Prob(F-statistic)	0.000000

جدول ۳

Dependent Variable: LOG(G/POP)
Method: Least Squares
Date: 11/03/04 Time: 18:37
Sample: 1974 1995
Included observations: 22

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	15.84118	6.839365	2.316177	0.0341
LOG(Y)	1.014076	0.285686	3.549621	0.0027
LOG(PG)	-0.242374	0.040030	-6.054849	0.0000
LOG(PNC)	0.330168	0.181117	1.822947	0.0871
LOG(PUC)	-0.055374	0.050774	-1.090602	0.2916
YEAR	-0.012617	0.003774	-3.343525	0.0041

R-squared	0.920662	Mean dependent var	0.081938
Adjusted R-squared	0.895869	S.D. dependent var	0.052899
S.E. of regression	0.017070	Akaike info criterion	-5.075987
Sum squared resid	0.004662	Schwarz criterion	-4.778430
Log likelihood	61.83585	F-statistic	37.13387
Durbin-Watson stat	1.422049	Prob(F-statistic)	0.000000

نسبت F برای آزمون برابری ضرایب در دو دوره مزبور بصورت زیر است. در مدل مقید ضرایب را یکسان فرض می کنیم.

$$F[6, 24]=\frac{[0.0252188-(0.000652+0.004662)]/6}{(0.000652+0.004662)/(14+22-12)}=14.968$$

مقدار بحرانی از جدول F برابر با ۲/۵۱ است. و لذا فرضیه برابری ضرایب در این دوره رد میشود.

۶-۵-۱ عرض از مبداهای مختلف:

ممکن است محققی بگوید پس از شوک ۱۹۷۳، کشورهای غربی مصرف بنزین را باندازه سهم ثابتی کاهش داده اند، اما سایر روابط در بازار مانند کشش درآمدی بدون تغییر باقی مانده است. این اتفاق مدل لگاریتمی را در مقدار ثابت کاهش می دهد.

لذا مدل بدون قید، مانند مدل قبلی اما با ضرایب متفاوت در دو دوره می باشد، در صورتیکه مدل مقید، یا Pooled در اینجا با عرض از مبدا متفاوت است. ماتریس رگرسورها بصورت زیر خواهد بود.

$$X_u=\begin{bmatrix} i & o & \underset{\leftarrow 73}{W} & o \\ 0 & i & o & \underset{\rightarrow 73}{W} \end{bmatrix}$$ مدل

$$X_R=\begin{bmatrix} i & o & \underset{\leftarrow 73}{W} \\ o & i & \underset{\rightarrow 73}{W} \end{bmatrix}$$ مدل

دو ستون اول X متغیرهای مجازی هستند. لذا سه رگرسیون انجام میشود که نتایج آنها در جداول ۴-۵-۶ آمده است.

جدول ۴

Dependent Variable: LOG(G/POP)
Method: Least Squares
Date: 11/03/04 Time: 19:08
Sample: 1960 1995
Included observations: 36

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D1	16.65782	8.654246	1.924815	0.0641
D2	16.73513	8.651836	1.934287	0.0629
LOG(Y)	1.838167	0.167258	10.98998	0.0000
LOG(PG)	-0.178005	0.033508	-5.312348	0.0000
LOG(PNC)	0.209842	0.129267	1.623331	0.1153
LOG(PUC)	-0.128132	0.060721	-2.110154	0.0436
YEAR	-0.016862	0.005105	-3.303111	0.0025

R-squared	0.978142	Mean dependent var	-0.003709
Adjusted R-squared	0.973620	S.D. dependent var	0.151691
S.E. of regression	0.024638	Akaike info criterion	-4.396414
Sum squared resid	0.017603	Schwarz criterion	-4.088508
Log likelihood	86.13546	F-statistic	216.2908
Durbin-Watson stat	0.977085	Prob(F-statistic)	0.000000

جدول ۵

Dependent Variable: LOG(G/POP)
Method: Least Squares
Date: 11/03/04 Time: 19:10
Sample: 1960 1973
Included observations: 14

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D1	-55.78632	18.04991	-3.090670	0.0149
LOG(Y)	0.423995	0.351778	1.205291	0.2625
LOG(PG)	0.094547	0.150511	0.628172	0.5474
LOG(PNC)	0.583896	0.131493	4.440510	0.0022
LOG(PUC)	-0.334619	0.092326	-3.624299	0.0067
YEAR	0.026366	0.010690	2.466382	0.0389

R-squared	0.998033	Mean dependent var	-0.138297
Adjusted R-squared	0.996804	S.D. dependent var	0.159715
S.E. of regression	0.009030	Akaike info criterion	-6.279088
Sum squared resid	0.000652	Schwarz criterion	-6.005206
Log likelihood	49.95362	F-statistic	811.8469
Durbin-Watson stat	2.391035	Prob(F-statistic)	0.000000

جدول ۶

Dependent Variable: LOG(G/POP)
Method: Least Squares
Date: 11/03/04 Time: 19:11
Sample: 1974 1995
Included observations: 22

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D2	15.84118	6.839365	2.316177	0.0341
LOG(Y)	1.014076	0.285686	3.549621	0.0027
LOG(PG)	-0.242374	0.040030	-6.054849	0.0000
LOG(PNC)	0.330168	0.181117	1.822947	0.0871
LOG(PUC)	-0.055374	0.050774	-1.090602	0.2916
YEAR	-0.012617	0.003774	-3.343525	0.0041

R-squared	0.920662	Mean dependent var	0.081938
Adjusted R-squared	0.895869	S.D. dependent var	0.052899
S.E. of regression	0.017070	Akaike info criterion	-5.075987
Sum squared resid	0.004662	Schwarz criterion	-4.778430
Log likelihood	61.83585	F-statistic	37.13387
Durbin-Watson stat	1.422049	Prob(F-statistic)	0.000000

بر اساس مقادیر مجموع مجذورات پسماندهای سه رگرسیون بالا F محاسبه شده بصورت زیر خواهد بود:

$$F=[5,24]=\frac{(0.017603-0.000652-0.004662)/5}{(0.0006522+(0.004662)/(14+22-12)}=11.099$$

مقدار بحرانی F از جدول برابر با ۲/۶۲ است، لذا این فرضیه نیز رد میشود.

داده‌های آماری می‌گوید که برای دو دوره مختلف بطور سیستماتیک ضرایب متفاوتند و فقط تغییر در مقدار ثابت نیست.

۶-۷-۲ تغییردر برخی از ضرایب:

نتایج جدول قبل نشان میدهد که عرض از مبداء کشش قیمتی و درآمدی بیش از کششهای متقاطع و ضریب روند زمان تغییر کرده‌اند. آزمونChow برای این موارد از محدودیت، تقریبا شبیه مورد تغییر در عرض از مبداهاست.

فرض کنیم ماتریس Z متغیرهای قیمت بنزین و درآمد را نشان دهد که ممکن است تغییر کرده باشند و ماتریس W قیمت اتومبیل و روند زمان باشد که اعتقاد بر این است ضرایب آنها ثابت باقی مانده است. لذا ماتریس رگرسیون مقید بصورت زیر است:

$$X=\begin{bmatrix} i_{pre} & Z_{pre} & o & o & W_{pre} \\ o & o & i_{post} & Z_{post} & W_{pre} \end{bmatrix}$$

مانند قبل بردار ضرایب در مدل غیر مقید از ترکیب دو مدل جدا از هم به دست می آید نسبت F برای آزمون فوق عبارتست از:

$$F[3/24]=\frac{(0/0080200 - 0/00065227 - 0/0046621)/3}{(0/00065227 + 0/0046621)/(14+22-12)}=4/086$$

مقدار بحرانی از جدول F، برابر با ۳/۰۱ است لذا این فرضیه نیز رد میشود.

ملاحظه میشود F در این مدل خیلی کوچکتر از دو F قبلی است. می توانیم بگوئیم دلیل این امر آنست که تفاوت در مدلها در دو دوره مختلف توسط تغییرات در مقدار ثابت، قیمت و درآمد قابل توضیح است.

فصل هفتم

فصل هفتم

۱-۷ مدل رگرسیون خطی تعمیم یافته

در مدل‌های قبلی داشتیم:

$$y = X\beta + u$$
$$E(u) = 0$$
$$E(u'u) = \sigma^2 I$$
$rank(X) = K \leq T$ و غیرتصادفی X

مواردی وجود دارد که ماتریس واریانس - کوواریانس جمله اخلال بصورت فوق نیست:

مثال(۱): در داده‌های مقطعی[۲۵]

$$C_i = a + by_i + u_i$$

وقتی Y بالا می‌رود واریانس نیز افزایش می‌یابد لـذا σ^2 ثابت نیسـت. کـه بـه آن واریـانس ناهمسـانی[۲۶] می‌گویند. چون میزان پراکندگی مصرف خانواده‌های پردرآمـد نسـبت بـه میـانگین مصـرف بیشـتر از میـزان پراکندگی مصرف خانواده‌های کم‌درآمد نسبت به میانگین مصرف جامعه است.

---

25 . Cross-Section Data
26 . Heteroscedastisity

لـذا:

$$E\left(u_i^2\right)=\sigma_i^2=\alpha y_i^2\left(\alpha>0\right)$$

$$E\left(u'u\right)=\alpha\begin{bmatrix} y_1^2 & 0 & \cdots & 0 \\ 0 & y_2^2 & & \vdots \\ \vdots & 0 & & \\ & \vdots & & \\ 0 & \ldots & & y_i^2 \end{bmatrix}$$

مثال (۲) خودهمبستگی[۲۷] یا هم‌بستگی پیاپی در داده‌های زمانی

$$C_t=a+by_t+u_t$$
$$u_t=\rho u_{t-1}+\varepsilon_t \qquad |\rho|<1$$
$$E\left(\varepsilon_t\right)=0 \quad , \quad E\left(\varepsilon_t^2\right)=\sigma_\varepsilon^2 \quad , \quad E\left(\varepsilon_t\varepsilon_s\right)=0 \quad t\neq s$$

$$E\left(uu'\right)=\frac{\sigma_3^2}{1-\rho^2}\begin{bmatrix} 1 & \rho & \rho^2 & \ldots & \rho^{T-1} \\ \rho & 1 & \rho & \ldots & \rho^{T-2} \\ \rho^2 & \rho & 1 & \ldots & \rho^{T-3} \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

لذا بطورکلی ممکن است با مدلهایی مواجه شویم که ماتریس واریانس-کوواریانس جمله اخـلال بصـورت زیر باشد:

$$E\left(uu'\right)=\sigma^2\Omega$$

بطوریکـه Ω یـک مـاتریس مثبـت معـین مـی‌باشـد و مـی‌توانـد معلـوم یـا مجهـول باشـد. و فرض کنیم $\Omega^{-1}=P'P \Rightarrow \Omega=P^{-1}P^{-1'}$ P غیرمنفرد است.

۷-۲ روش حداقل مربعات تعمیم‌یافتـه[۲۸] (GLS)

مدل اصلی را با استفاده از ماتریس P تبدیل می‌کنیم:

$$y=X\beta+u$$
$$Py=PX\beta+Pu$$
$$y_*=X_*\beta+u_*$$

27 Autorgressive
28 . Generalized Least Square

فروض کلاسیک را در مدل تبدیل یافته شده بررسی می‌کنیم:

$1) X_*$ غیرتصادفی

چون P غیرمنفرد است.

$$2) rank(X_*) = rank(PX) = rank(X) = K$$

$$3) E(u_*) = E(Pu) = PE(u) = 0$$

$$4) E\left(u_* u_*'\right) = E(Puu'P') = PE(uu')P' = P\sigma^2\Omega P'$$

$$= \sigma^2 P\Omega P' = \sigma^2 PP^{-1}P^{-1'}P' = \sigma^2 I$$

لـذا تخمین‌زننده GLS را زمانیکه ماتریس واریانس کوواریانس معلوم است می‌توان بصورت زیـر تعریـف نمود:

$$\tilde{\beta} = (X_*'X_*)^{-1} X_*'y_*$$

$$\tilde{\beta} = (X'P'PX)^{-1} X'P'Py$$

$$\tilde{\beta} = (X'\Omega^{-1}X)^{-1} X'\Omega^{-1}y$$

مثال: مدل مصرف را در داده‌های مقطعی در نظر بگیریم:

$$C_i = a + by_i + u_i \qquad E(u_i) = 0 \qquad , \quad E(u_i^2) = \sigma_i^2 \qquad i = 1, \ldots, N$$

$$E(u'u) = \begin{bmatrix} \sigma_1^2 & 0 & \ldots & 0 \\ & \sigma_2^2 & \ldots & 0 \\ & & & \vdots \\ & \ldots & & \sigma_N^2 \end{bmatrix} = \Omega, \quad \Omega^{-1} = \begin{bmatrix} \frac{1}{\sigma_1^2} & 0 & \ldots & 0 \\ & \frac{1}{\sigma_1^2} & \ldots & 0 \\ & \ldots & & \frac{1}{\sigma_N^2} \end{bmatrix} = P'P$$

$$P = \begin{bmatrix} \frac{1}{\sigma_1} & 0 & \ldots & 0 \\ & \frac{1}{\sigma_2} & \ldots & 0 \\ 0 & \ldots & & \frac{1}{\sigma_N} \end{bmatrix}$$

$$PC = P\begin{bmatrix} 1 & y_i \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} + Pu$$

$$P\begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_N \end{bmatrix} = P\begin{bmatrix} 1 & y_1 \\ 1 & \vdots \\ 1 & \\ 1 & \\ 1 & y_N \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} + P\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_N \end{bmatrix}$$

و یا :

$$\frac{C_i}{\sigma_i} = \frac{\alpha}{\sigma_i} + b\frac{y_i}{\sigma_i} + \frac{u_i}{\sigma_i}$$

$$C_* = a + by_* + u_* \qquad E(u_*^2) = 1$$

قضیه ۱) $\tilde{\beta}_{GLS}$ یک تخمین‌زننده بدون تورش است.

$$\tilde{\beta} = (X'\Omega^{-1}X)^{-1} X'\Omega^{-1}y$$

$$\tilde{\beta} = (X'\Omega^{-1}X)^{-1} X'\Omega^{-1}X\beta + (X'\Omega^{-1}X)^{-1} X'\Omega^{-1}u$$

$$E(\tilde{\beta}) = \beta + E\left[(X'\Omega^{-1}X)^{-1} X'\Omega^{-1}u\right]$$

$$= \beta + (X'\Omega^{-1}X)^{-1} X'\Omega^{-1}E(u)$$

$$E(\tilde{\beta}) = \beta$$

ماتریس واریانس-کوواریانس $\tilde{\beta}$ عبارتست از:

$$\tilde{\beta} - \beta = (X'\Omega^{-1}X)^{-1} X'\Omega^{-1}u$$

$$Vc(\tilde{\beta}) = E(\tilde{\beta} - \beta)(\tilde{\beta} - \beta)' = E\left[(X'\Omega^{-1}X)^{-1} X'\Omega^{-1}uu'\Omega^{-1}X(X'\Omega^{-1}X)^{-1}\right]$$

$$= (X'\Omega^{-1}X)^{-1} X'\Omega^{-1}E(uu')\Omega^{-1}X(X'\Omega^{-1}X)^{-1}$$

چون $E(uu') = \sigma^2\Omega$ است پس:

$$Vc(\tilde{\beta}) = \sigma^2(X'\Omega^{-1}X)^{-1}$$

۳-۷ تئوری ایتکنز [۲۹]

این قضیه بیان می‌کند که $\tilde{\beta}_{GLS}$ تخمین‌زننده‌ای خطی بدون تورش با کمترین واریانس است. فرض کنیم فروض ۱ تا ۳ برقرار است و هم‌چنین فرض کنیم دو تخمین‌زننده $\left(\check{\beta},\tilde{\beta}\right)$ برای β وجود دارد.

نشان خواهیم داد که $\tilde{\beta}$ یک تخمین‌زننده $BLUE$ می‌باشد.

اثبات: اگر $\check{\beta}=Cy$ در نظر بگیریم بطوریکه $\mathrm{C}=\left(\mathrm{X}'\Omega^{-1}\mathrm{X}\right)^{-1}\mathrm{X}'\Omega^{-1}+\mathrm{D}$ بدون تورش بودن $\check{\beta}$ نتیجه گرفته می‌شود.

$$\mathrm{E}\left(\check{\beta}\right)=\beta+\mathrm{E}(\mathrm{Dy})=\beta+\mathrm{DE}(\mathrm{X}\beta+\mathrm{u})=\beta+\mathrm{DX}\beta=\beta(\mathrm{I}+\mathrm{DX})$$

اگر $DX=0$ باشد $\check{\beta}$ بدون تورش است.

$$\check{\beta}=\tilde{\beta}+Dy$$

$$Vc\left(\check{\beta}\right)=E\left(\check{\beta}-\beta\right)\left(\check{\beta}-\beta\right)'$$

$$Vc\left(\check{\beta}\right)=E\left(\tilde{\beta}-\beta+Dy\right)\left(\tilde{\beta}-\beta+Dy\right)'$$

$$Dy=DX\beta+Du, DX=0\Rightarrow Dy=Du$$

$$Vc\left(\check{\beta}\right)=E\left(\left(\tilde{\beta}-\beta+Du\right)\left(\tilde{\beta}+\beta+Du\right)'\right)$$

$$=E\left(\tilde{\beta}-\beta\right)\left(\tilde{\beta}-\beta\right)'+E\left(\tilde{\beta}-\beta\right)u'D'+EDu\left(\tilde{\beta}-\beta\right)'+E(Duu'D') \qquad (1)$$

$$=Vc\left(\tilde{\beta}\right)+\sigma^{2}D\Omega D'$$

جمله دوم رابطه (۱) برابر است با:

$$E\left(X'\Omega^{-1}X\right)^{-1}X'\Omega^{-1}uu'D'$$

$$=\left(X'\Omega^{-1}X\right)^{-1}X'\Omega^{-1}\sigma^{2}\Omega D'$$

$$=\sigma^{2}\left(X'\Omega^{-1}X\right)^{-1}X'D'=0$$

چون $DX=0$

بهمین ترتیب برای جمله سوم نشان می‌دهیم، که برابر صفر است. لذا

$$Vc\left(\check{\beta}\right)=Vc\left(\tilde{\beta}\right)+\sigma^{2}D\Omega D'$$

$$Vc\left(\check{\beta}\right)-Vc\left(\tilde{\beta}\right)>0$$

29 .Aitken's-Theorem

همه عناصر روی قطر $Vc\left(\check{\beta}\right)$ بزرگتر از عناصر متناظر بر روی قطر ماتریس $Vc\left(\tilde{\beta}\right)$ می‌باشند. پس $\tilde{\beta}$ دارای کمترین واریانس است و قضیه ثابت می‌گردد.

فصل هشتم

فصل هشتم

۸-۱ مدل رگرسیونی ظاهراً نامرتبط[۳۰]

فرض کنید دو معادله رگرسیون خطی داریم:

$$y_1 = x_1 \beta_1 + u_1 \qquad Eu_1 = 0 \quad , \quad Eu_1 u_1' = \sigma_1^2 I$$
$$y_2 = x_2 \beta_2 + u_2 \qquad Eu_2 = 0 \quad , \quad Eu_2 u_2' = \sigma_2^2 I$$

ولی $Eu_1 u_2' \neq 0$

$$E \begin{bmatrix} u_{11} \\ u_{21} \\ \vdots \\ u_{T_1} \end{bmatrix} \begin{bmatrix} u_{12} \cdots u_{T2} \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \, I_T \end{bmatrix}$$

$$y_{.1} = X_i \beta_i + u_i \qquad i = 1, \ldots, M$$

$$y_{.i} = \begin{bmatrix} y_{1i} \\ \vdots \\ y_{T_i} \end{bmatrix} \qquad X_{iTK.i} \qquad \beta_{iki.1} \qquad u_{.i} = \begin{bmatrix} u_{1i} \\ u_{2i} \\ \vdots \\ u_{T_i} \end{bmatrix}$$

فروض زیر را در نظر بگیریم:

i) $E(u_{.i}) = 0 \qquad \text{for} \qquad i = 1, \ldots, M$

ii) $Vc(u_{.i} u_{.j}') = \sigma_{ij} I_T$

برای مثال:

---

30 - Seemingly Unrelated Regression (SUR).

اگر $t = s$ باشد داریم: $E\,u_{1i}\,u_{1j} = \sigma_{ij}$

اگر $t \neq s$ باشد داریم: $E u_{ti}\ u_{sj} = 0$

هر دو معادله دارای یک ماتریس واریانس کوواریانس، بصورت $\sigma_{ij}\,I_T$ است.

iii) X_i تصادفی نیستند

$v)\ \text{rank}\,(X_i) = k_i$

$y = X\beta + u$

مدل را بصورت ماتریسی می‌نویسیم:

$$\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_M \end{bmatrix}_{\sum ki} ,\quad y = \begin{bmatrix} y_{.1} \\ y_{.2} \\ \vdots \\ y_{.M} \end{bmatrix}_{TM.1} ,\quad u = \begin{bmatrix} u \\ u_{.2} \\ \vdots \\ u_{.M} \end{bmatrix} ,\quad X = \begin{bmatrix} X_1 & 0 & \cdots & 0 \\ & & & \vdots \\ 0 & X_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & X_M \end{bmatrix}_{TM.\sum ki}$$

$$E(u) = 0$$

$$Euu' = E\begin{bmatrix} u_{.1}u'_{.1} & u_{.1}u'_{.2} & \cdots \\ u_{.2}u'_{.1} & u_{.2}u'_{.2} & \cdots \\ \vdots & \vdots & \\ \cdots & u_{.i}u'_{.j} & \cdots \end{bmatrix} = \begin{bmatrix} \sigma_{11} I_T & \sigma_{12} I_T & \cdots & \sigma_{1M} I_T \\ \sigma_{21} I_T & \sigma_{22} I_T & \cdots & \sigma_{2M} I_T \\ \vdots & \vdots & & \vdots \\ \sigma_{M1} I_T & \sigma_{M2} I_T & \cdots & \sigma_{MM} I_T \end{bmatrix}$$

$$= \sum \otimes I_T = \Omega$$

$$\hat{\beta}_{OLS} = (X'X)^{-1}X'y = \hat{\beta}_{i,OLS} = (X'_i X_i)^{-1} X'_i\, y_{.i}$$

$$\hat{\beta}_{OLS} = \begin{bmatrix} \hat{\beta}_{i,OLS} \\ \vdots \\ \hat{\beta}_M \end{bmatrix} = \begin{bmatrix} (X'_1 X_1)^{-1} X'_1 y_{.1} \\ \vdots \\ (X'_M X_M)^{-1} X'_M y_{.M} \end{bmatrix}$$

شرایط و نتایج معادلات انفرادی برای همه معادلات نیز برقرار است در اینجا OLS کارآ نیست و بهترین نخواهد بود. GLS از کارآیی بالاتری برخوردار است.

$\tilde{\beta}_{GLS}$ وقتی Ω معلوم است

$$\tilde{\beta}_{GLS} = (X'\ \Omega^{-1} X_1)^{-1} X'\ \Omega^{-1} y$$

$$\Omega = \sum \otimes I_T \quad \Rightarrow \quad \Omega^{-1} = \sum^{-1} \otimes I_T$$

$$\sum = (\sigma_{ij})\ ,\ \sum^{-1} = (\sigma^{ij}) \qquad i, j = 1, \ldots, M$$

$$\tilde{\beta}_{GLS} = \left[X'\ (\sum^{-1} \otimes I_T)\ X\right]^{-1}\ X'\ (\sum^{-1} \otimes I_T)\ y$$

$$\tilde{\beta}_{GLS} = \begin{bmatrix} \sigma^{11} X_1' X_1 & \sigma^{12} X_1' X_2 & \cdots & \sigma^{1M} X_1' X_M \\ \vdots & \vdots & & \vdots \\ \sigma^{M1} X_M' X_1 & \cdots & & \sigma^{MM} X_M' X_M \end{bmatrix}^{-1} \begin{bmatrix} \sum_{j=1}^{M} & \sigma^{1j} & X_1' y_{.j} \\ \vdots & \vdots & \vdots \\ \sum_{j=1}^{M} & \sigma^{Mj} & X_M' y_{.j} \end{bmatrix}$$

$$Vc(\tilde{\beta}_{GLS}) = (X' \Omega^{-1} X)^{-1} = \left[X' \left(\sum{}^{-1} \otimes I_T\right) X\right]^{-1}$$

قضیه: دو حالت وجود دارد که $\tilde{\beta}_{GLS} = \tilde{\beta}_{OLS}$ است.

الف) اگر $i \neq j$ آنگاه $\sigma_{ij} = 0$

ب) $X_1 = X_2 = \dots = \bar{\bar{X}}$

اثبات

الف) وقتی جملات اخلال معادلات مختلف با یکدیگر ارتباط ندارند و امکان همبستگی همزمان[31] را نادیده بگیریم و مدل را تخمین بزنیم چیزی را از دست نخواهیم داد:

$$\sigma_{ij} = 0 \quad \Rightarrow \quad (X' \Omega^{-1} X)^{-1} = \begin{bmatrix} \sigma^{11} X_1' X_1 & \cdots & \\ \vdots & \sigma^{ii} X_2' X_2 & 0 \\ 0 & \cdots & \sigma^{MM} X_M' X_M \end{bmatrix}^{-1}$$

و

$$(X' \Omega^{-1} y) = \begin{bmatrix} \sigma^{11} X_1' y_{.j} \\ . \\ . \\ \sigma^{MM} X_M' y_{.j} \end{bmatrix}$$

$$\tilde{\beta}_{GLS} = (X' \Omega^{-1} X)^{-1} X' \Omega^{-1} y =$$

$$= \begin{bmatrix} \sigma_{11}(X_1' X_1)^{-1} & \cdots & \\ \vdots & \sigma_{ii}(X_2' X_2)^{-1} & 0 \\ 0 & \cdots & \sigma_{MM}(X_M' X_M)^{-1} \end{bmatrix} \begin{bmatrix} \sigma^{11} X_1' y_{.1} \\ . \\ . \\ . \\ \sigma^{MM} X'_M y_{.M} \end{bmatrix}$$

$$\therefore \tilde{\beta}_{GLS} = \begin{bmatrix} (X_1' X_1)^{-1} X_1' y_{.1} \\ \vdots \\ (X_M' X_M)^{-1} X_M' y_{.M} \end{bmatrix} = \tilde{\beta}_{OLS} \qquad \text{چون} \quad \sigma^{ii}.\sigma_{ii} = I$$

31 - Contemporaneously Correlation.

ب) چنانچه $X_1 = X_2 = \ldots = \bar{\bar{X}}$

$$X = I_M \otimes \bar{\bar{X}}$$

$$\tilde{\beta}_{GLS} = \left[X' \left(\sum^{-1} \otimes I_T\right) X\right]^{-1} X' \left(\sum^{-1} \otimes I_T\right) y$$

$$X' = I_M \otimes \bar{\bar{X}}'$$

$$\tilde{\beta}_{GLS} = \left[(I_M \otimes \bar{\bar{X}})' \left(\sum^{-1} \otimes I_T\right) (I_M \otimes \bar{\bar{X}})\right]^{-1} (I_M \otimes \bar{\bar{X}})' \left(\sum^{-1} \otimes I_T\right) y$$

$$\tilde{\beta}_{GLS} = \left[\left(\sum^{-1} \otimes \bar{\bar{X}}'\right) (I_M \otimes \bar{\bar{X}})\right]^{-1} \left(\sum^{-1} \otimes \bar{\bar{X}}'\right) y$$

$$= \left[\left(\sum^{-1} \otimes \bar{\bar{X}}' \bar{\bar{X}}\right)\right]^{-1} \left(\sum^{-1} \otimes \bar{\bar{X}}'\right) y$$

$$= \left(\sum \otimes (\bar{\bar{X}}' \bar{\bar{X}})^{-1}\right) \left(\sum^{-1} \otimes \bar{\bar{X}}'\right) y$$

$$= I_M \otimes (\bar{\bar{X}}' \bar{\bar{X}})^{-1} \bar{\bar{X}}' y$$

$$= \begin{bmatrix} (\bar{\bar{X}}' \bar{\bar{X}})^{-1} \bar{\bar{X}}' & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & & \\ 0 & \cdots & 0 & (\bar{\bar{X}}' \bar{\bar{X}})^{-1} \bar{\bar{X}}' \end{bmatrix} \begin{bmatrix} y_{.1} \\ \vdots \\ y_{.M} \end{bmatrix} = \beta_{OLS}$$

اگر Σ معلوم نباشد باید $\hat{\Sigma}$ را بکار برد. در آنصورت با Feasible GLS مواجه خواهیم شد.

$$\hat{\beta}_{FFLS} = \left[X'(\hat{\Sigma}^{-1} \otimes I_T) X\right]^{-1} X'(\hat{\Sigma}^{-1} \otimes I_T) y \quad \hat{\Omega} = \hat{\Sigma} \otimes I_T$$

که $\hat{\Omega}$ تخمین‌زننده سازگاری از Ω خواهد بود.

فصل نهم

۱-۹ تخمین به روش حداکثر نمودن تابع درست‌نمایی

روش تخمین حداکثر نمودن تابع درست‌نمایی روشی است که در مسایل گوناگونی کاربرد دارد. ولی در مدل رگرسیون خطی چنانچه توزیع جمله اخلال نرمال باشد تخمین‌زننده‌هایی مشابه روش حداقل مربعات معمولی بدست می‌آوریم.

مدل زیر را در نظر بگیریم:

$$y_i = \alpha + \beta x_i + u_i$$

بطوریکه:

$$u_i \sim IN(0, \sigma^2)$$

چون u دارای توزیع نرمال است و بطور مستقل توزیع شده است لذا، y_i بصورت نرمال و مستقل با میانگین $\alpha + \beta x_i$ و واریانس σ^2 توزیع شده است. تابع توزیع مشترک yiها (راست‌نمایی) برروی کل مشاهدات بصورت زیر است:

$$f(y_1, \ldots, y_n) = \prod_{i=1}^{n} \left(\frac{1}{2\pi\sigma^2} \right)^{\frac{1}{2}} \exp\left[-\frac{1}{2\sigma^2} (y_i - \alpha - \beta X_i)^2 \right]$$

این تابع وقتی بصورت تابعی از پارامترهای $(\alpha, \beta, \sigma^2)$ در نظر گرفته شود، تابع درستنمایی[۳۲] نامیده می‌شود و بصورت $L(\alpha, \beta, \sigma^2)$ نشان داده می‌شود. روش تخمین حداکثر درستنمایی (ML) پیشنهاد می‌کند که، مقادیر پارامترها را طوری انتخاب کنیم که تابع درستنمایی را حداکثر می‌نماید.

---

32 - Likelihood Function.

برای سادگی از تابع درستنمایی، لگاریتم می‌گیریم و چون تبدیل خطی است، نتایج به‌دست آمده نقاط مشابهی را بدست می‌دهد.

برای مدل رگرسیون خطی داریم:

$$\mathrm{Log\,L} = \sum_{i=1}^{n}\left[-\frac{1}{2}\,\mathrm{Log}(2\pi\sigma^2) - \frac{1}{2\sigma^2}\,(y_i - \alpha - \beta X_i)^2\right]$$
$$= c - \frac{n}{2}\,\mathrm{Log}\,\sigma^2 - \frac{Q}{2\sigma^2}$$

بطوریکه $c = \left(\frac{n}{2}\right)\mathrm{Log}\,(2\pi)$ مقدار ثابتی است و با پارامترها درگیر نمی‌باشد. از طرفی $Q = \Sigma\,(y_i - \alpha - \beta X_i)^2$ می‌باشد.

تابع $\mathrm{Log\,L}$ را ابتدا نسبت به β و α و سپس نسبت به σ^2 حداکثر می‌کنیم، توجه داریم که فقط جمله سوم تابع فوق درگیر پارامتر β و α می‌باشد، لذا حداکثر نمودن این تابع مانند مینیمم نمودن Q می‌باشد (چون این جمله دارای علامت منفی است). در نتیجه تخمین‌زننده‌های ML برای β و α عیناً برابر برآورد کننده‌های $\hat{\beta}$ و $\hat{\alpha}$ به روش حداقل مربعات معمولی (OLS) می‌آید.

چنانچه $\hat{\beta}$ و $\hat{\alpha}$ را در تابع درستنمایی فوق با β و α جایگزین کنیم، تابع ما فقط σ^2 تبعیت می‌کند. حال داریم:

$$LogL(\sigma^2) = Constant - \frac{n}{2}Log\sigma^2 - \frac{\hat{Q}}{2\sigma^2}$$

بطوریکه $\hat{Q} = \sum (y_i - \hat{\alpha} - \hat{\beta}x_i)^2$ چیزی جزمجموع مجذورات پسماندها نیست.

چنانچه ازآن نسبت به σ^2 مشتق گرفته وآنرا برابر صفر قرار دهیم داریم:

$$-\frac{n}{2\sigma^2} + \frac{\hat{Q}}{2\sigma^4} = 0$$

از رابطه فوق برآورده کننده ML برای σ^2 بصورت زیرخواهد بود.

$$\tilde{\sigma}^2 = \frac{\hat{Q}}{n} = \frac{\mathrm{RSS}}{n}$$

توجه کنیم این تخمین زننده با تخمین‌زننده OLS، بدون تورش برای σ^2 که بصورت $\frac{\mathrm{RSS}}{n-2}$ بود تفاوت دارد. چنانچه n افزایش یابد این دو روش به هم نزدیک خواهندشد. ML یک روش تخمین درنمونه‌های بزرگ است.

چنانچه در رابطه $LogL\,(\sigma^2)$، $\hat{\sigma}^2 = \frac{\hat{Q}}{n}$ را قرار دهیم، مقدار حداکثر تابع لگاریتم درست نمایی بصورت زیر بدست خواهد آمد:

$$\underset{\alpha\,\beta\,\sigma^2}{Max}\;LogL = c - \frac{n}{2}Log\frac{\hat{Q}}{n} - \frac{n}{2}$$
$$= c - \frac{n}{2}\mathrm{Log}\hat{Q} + \frac{n}{2}\mathrm{Log}\,n - \frac{n}{2}$$

این رابطه برای بدست آوردن آزمون نمودن β و α و σ^2 با استفاده از روش ML مفید خواهد بود.
چون ما حجم نمونه یعنی n را تغییر نمی‌دهیم، صرفاً می‌نویسیم:

$$Max\ LogL = (Constant) - \frac{n}{2} Log\hat{Q}$$

$$Max\ L = Constant.(\hat{Q})^{\frac{-n}{2}} = Constant(RSS)^{\frac{-n}{2}}$$

حال چنانچه در نظر داشته باشیم مدل رگرسیون چند متغیره را به روش حداکثر درستنمایی بدست آوریم بصورت زیر عمل می‌کنیم.

$$Y = X\beta + U$$

بطوریکه: $U \sim N(0, \sigma^2 I)$

چون: $f(u_i)\dfrac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{u^2}{2\sigma^2}}$

از طرفی بردار Y نیز بصورت زیر توزیع شده است:

$$Y \sim N(X\beta, \sigma^2 I)$$

و قانون توزیع Y_iکه بصورت i.i.d توزیع شده‌اند عبارتست از:

$$f(y_i) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(y_i - x\beta)^2\right)$$

تابع درستنمایی y عبارتست از[۳۳]:

$$L(y; x\beta, \sigma^2) = f(y_1, y_2, \ldots, y_n; x\beta, \sigma^2)$$

$$= \prod_{i=1}^{T} f(y_i)$$

$$= (2\pi\sigma^2)^{\frac{-T}{2}} \exp\left[-\frac{1}{2\sigma^2}\Sigma(y_i - x_i\beta)^2\right]$$

$$L^* = Log\,L = -\frac{T}{2} Log\,2\pi - \frac{-T}{2} Log\,\sigma^2 - \frac{1}{2\sigma^2}(y_i - x\beta)'(y - x\beta)$$

$$\frac{\partial L^*}{\partial \beta} = \frac{1}{\sigma^2} x'(y - x\beta) = 0 \quad \Rightarrow \quad \hat{\beta} = (x'x)^{-1} x'y$$

$$\frac{\partial L^*}{\partial \sigma^2} = -\frac{T}{2\sigma^2} + \frac{1}{2\sigma^4}(y - x\beta)'(y - x\beta) = 0$$

$$\hat{\sigma}^2_{ML} = \frac{\hat{U}'\hat{U}}{T} = \frac{Y'MY}{T}$$

33 - *if* $X \sim N(\mu, \Sigma)$ *then* $f(X_1, X_2 \ldots X_n) = \dfrac{1}{\sqrt{2\pi|\Sigma|}} e^{-\frac{1}{2}(X_i - \mu)'\Sigma^{-1}(X_i - \mu)}$

$$\frac{T}{T-K}\hat{\sigma}^2_{ML}=\hat{\sigma}^2_{OLS}$$

در این روش تخمین پارامترها بنحوی انتخاب می‌شوند که احتمال بدست آمدن نمونه را ماکزیمم می‌کنند.

۹-۲ مقایسه واریانس $\hat{\sigma}^2_{OLS}$ و $\tilde{\sigma}^2_{ML}$

تخمین ML برای σ^2 فرم درجه دوم و تابعی از متغیر تصادفی با میانگین زیر است:

$$E(\tilde{\sigma}^2_{ML})=\frac{1}{T}E[\hat{u}'\hat{u}]\ \frac{1}{T}[u'Mu]=\frac{1}{T}\sigma^2(T-K)$$

و لذا دارای تورش است که تورش آن برابر است با $\frac{K}{T}\sigma^2$.

اگر چنانچه آن را ضرب $\frac{T}{T-K}$ تعدیل کنیم خواهیم داشت:

$$\frac{T}{T-K}.\ \tilde{\sigma}^2_{ML}=\hat{\sigma}^2_{OLS}$$

که تخمین زننده بدون تورش خواهد بود.

از طرفی $(T-K)\frac{\hat{\sigma}^2}{\sigma^2}$ یک فرم درجه دوم است که بصورت χ^2 با درجه آزادی T-K توزیع شده است. با استفاده از قوانین توزیع فوق می‌توانیم امید ریاضی و واریانس $\hat{\sigma}^2_{OLS}$ و $\tilde{\sigma}^2_{ML}$ را بدست آوریم.

$$E(\hat{\sigma}^2_{OLS})=\frac{\sigma^2}{T-K}\ E(\chi^2_{T-K})=\sigma^2$$

$$Var\ (\hat{\sigma}^2_{OLS})=E(\hat{\sigma}^2_{OLS}-\sigma^2)^2=\frac{\sigma^4}{(T-K)^2}\ Var\ (\chi^2_{T-K})=2\sigma^4/T-K$$

بهمین ترتیب میانگین و واریانس $\tilde{\sigma}^2_{ML}$ را بدست می‌آوریم:

$$E\ (\tilde{\sigma}^2{}_{ML})=\frac{\sigma^2}{T}\ E\ (\chi^2_{T-K})=\frac{T-K}{T}\ \sigma^2$$

$$Var\ \ (\tilde{\sigma}^2_{ML})=E\left\{\tilde{\sigma}^2_{ML}-E\left(\tilde{\sigma}^2_{ML}\right)\right\}^2=\frac{\sigma^4}{T^2}\ Var\ \ (\chi^2_{T-K})=\frac{2\sigma^4\ (T-K)}{T^2}$$

پس $\tilde{\sigma}^2_{ML}$ دارای تورش باندازه $-\left(\frac{K}{T}\right)\sigma^2$ و دارای واریانسی است که در نمونه‌های کوچک، کوچکتر از واریانس $\hat{\sigma}^2_{OLS}$ است. لذا محقق می‌تواند بین تورش و کوچکی واریانس انتخاب کند.

۹-۳ ماتریس اطلاعات[34]

ماتریس اطلاعات عبارتست از منهای امید ریاضی مشتق دوم تابع درستنمایی. یعنی اگر θ تخمین‌زننده θ, ML باشد و لگاریتم تابع درستنمایی $Log\ L(X;\theta)=Log\prod_{i=1}^{N}f(x_i;\theta)$

34 - Information Matrix.

آنگاه:

$$I(\theta)=-E\left[\frac{\partial^2 Log\, L(x;\theta)}{\partial\theta\,\partial\theta'}\right]$$

بطوریکه $(x_i;\theta)$ نسبت به مشتقات درجه اولش منظم[۳۵] باشد.

حد پایین رائو و کرامر[۳۶]

به معکوس ماتریس اطلاعات حد پایین رائو- کرامر گفته می‌شود. فرض کنید $\tilde{\theta}$ هر تخمین‌زننده بدون تورش با ماتریس واریانس – کوواریانس Σ باشد.
آنگاه ماتریس:

$$\Sigma - I^{-1}(\tilde{\theta}) > 0 \quad P.S.D$$

لذا اگر تخمین‌زننده‌ای دارای ماتریس واریانس- کوواریانس $I^{-1}(\tilde{\theta})$ باشد، کارآ است. ممکن است تخمین‌زننده‌ای کارآ باشد ولی حد پایین R.C را تامین نکند.
حال چنانچه بخواهیم ماتریس اطلاعات و حد پایین رائو- کرامر را برای مدل رگرسیون خطی مورد نظر بدست آوریم، ابتدا باید از شرایط اولیه ماکزیمم نمودن تابع درستنمایی مجدداً نسبت به بردار β و σ^2 مشتق بگیریم.

$$\frac{\partial^2 L^*}{\partial\beta\partial\beta'}=-\frac{1}{\sigma^2}X'X \qquad (۱)$$

$$\frac{\partial^2 L^*}{\partial\sigma^2\partial\sigma^2}=\frac{T}{2\sigma^4}-\frac{1}{\sigma^6}(Y-X\beta)'(Y-X\beta) \qquad (۲)$$

$$\frac{\partial^2 L^*}{\partial\beta\partial\sigma^2}=-\frac{1}{\sigma^4}X'(Y-X\beta)=\frac{\partial^2 L^*}{\partial\sigma^2\partial\beta} \qquad (۳)$$

برای تشکیل ماتریس اطلاعات لازمست از روابط ۱ تا ۳ منهای امید ریاضی بگیریم:

$$-E\left[\frac{\partial^2 L^*}{\partial\beta\partial\beta'}\right]=\frac{1}{\sigma^2}X'X$$

$$-E\left[\frac{\partial^2 L^*}{\partial\sigma^2\partial\sigma^2}\right]=\frac{T}{2\sigma^4}$$

$$-E\left[\frac{\partial^2 L^*}{\partial\beta\partial\sigma^2}\right]=0$$

و ماتریس اطلاعات عبارتست از:

۳۵- تابع منظم آنست که نسبت به مشتقات درجه اولش داراي اميد صفر باشد، يعني:

$$E\left[\frac{\partial Log\, L(x;\theta)}{\partial\theta_i}\right]=0 \quad i=1,\dots,N$$

36 - Rao- Cramer Lower Bond.

$$I(\beta,\sigma^2)=\begin{bmatrix}\frac{X'X}{\sigma^2} & 0\\ 0 & \frac{T}{2\sigma^4}\end{bmatrix}$$

۹-۳-۱ حد پایین رائو – کرامر

یعنی $I^{-1}(\beta,\sigma^2)$ عبارتست از:

$$I^{-1}(\beta,\sigma^2)=\begin{bmatrix}\sigma^2(X'X)^{-1} & 0\\ 0 & \frac{2\sigma^4}{T}\end{bmatrix}$$

لذا تخمین‌زننده‌های OLS برای ضریب β دارای حداقل واریانس می‌باشند و حد پایین رائو- کرامر را تامین می‌نمایند ولی همچنان که گفته شد:

$$Var(\tilde{\sigma}^2_{ML})<Var(\hat{\sigma}^2_{OLS})$$

$$\frac{2\sigma^4(T-K)}{T^2}<\frac{2\sigma^4}{T-K}$$

هر چند واریانس $\hat{\sigma}^2_{OLS}$ حد پایین رائو–کرامر را تامین نمی‌کند، ولی تخمین‌زننده بدون تورشی برای σ^2 که کمتر از $\frac{2\sigma^4}{T-K}$ باشد وجودندارد.

مثال: متغیر تصادفی x_i ، که دارای توزیع پوآسون می‌باشد را در نظر بگیرید:

$$f(x_i,\theta)=\frac{e^{-\theta}\theta^{x_i}}{x_i!}\qquad i=1,\dots,N$$

چنانچه xi ، i.i.d باشد، تخمین ML برای θ را بدست آورید.

حل: تابع توزیع مشترک

$$f(x_1,x_2,\dots,x_N;\theta)=L(x;\theta)$$

$$L(x;\theta)=\prod_{i=1}^{\tilde{}} f(x_i;\theta)$$

$$=\frac{e^{-N\theta}\theta^{\Sigma xi}}{\Pi x_i!}$$

$$L^*=Log\ L(x;\theta)=-N\theta+\Sigma XiLog\,\theta-Log\Pi x_i!$$

چنانچه از تابع فوق نسبت به θ مشتق بگیریم داریم:

$$\frac{dL^*}{d\theta}=-N+\frac{\Sigma X_i}{\hat{\theta}_{ML}}=0$$

$$\hat{\theta}_{ML}=\frac{\Sigma X_i}{N}=\bar{x}$$

$$\frac{d^2L^*}{d\theta^2}=-\frac{\Sigma X_i}{\hat{\theta}^2_{ML}}<0$$

پس تابع دارای ماکزیمم است.

حال اگر متغیر x مقادیر ۱، ۴، ۳، ۲، ۳، ۰، ۱، ۱، ۰ و ۵ را انتخاب کند $\hat{\theta}$ را بدست آورید.
(حل بعهده خودتان).

۹-۴ استفاده از آزمونهای LM, LR, WALD

چنانچه قصد داشته باشیم وجود رابطه خطی $R\beta=r$ را درون مدل رگرسیون خطی:

$$Y=X\beta+U$$

بررسی کنیم می‌توانیم از آزمون F که به آزمون WALD نیز معروف است و یا با استفاده از روش تابع درستنمایی و آزمون Likelihood Ratio و یا از آزمون Lagrange Multiplier استفاده نماییم.

فرض کنیم جمله اخلال دارای توزیع نرمال بصورت زیر باشد:

$$U \sim N(0,\sigma^2 I)$$

$$\Rightarrow \hat{\beta} \sim N(\beta,\sigma^2 (X'X)^{-1})$$

$$\Rightarrow R\hat{\beta} \sim N(R\beta,\sigma^2 R(X'X)^{-1}R')$$

چنانچه رابطه $R\beta=r$ صحیح باشد یعنی تحت شرایط H_O خواهیم داشت.

$$\frac{1}{\sigma^2}(R\hat{\beta}-r)'\left[R(X'X)^{-1}R'\right]^{-1}(R\hat{\beta}-r)\sim \chi^2_G$$

G تعداد قیدهای موجود برای آزمون و درجه آزادی چی دو خواهد شد. می‌دانیم که σ^2معمولاً در دسترس نیست ولذا بجای σ^2از $G.\sigma^2$ استفاده می‌کنیم که توزیع فوق به سمت $F_{G,T-K}$میل می‌کند که معروف بهWALD تست است و قبلاً هم بطور مفصل به آن پرداخته شد. ساختن آزمون فوق تحت هرعنوان با استفاده از روش تخمین OLS یاRLS می‌باشد.

هم چنین می‌توانیم با استفاده از تابع حداکثر راست نمایی آماره Likelihood Ratio را بصورت $-2\,Log\frac{L(\hat{\theta}_c)}{L(\hat{\theta})}$ تشکیل دهیم. که به سمت χ^2_G میل می‌نماید.

$\hat{\theta}_c$ تخمین مقید پارامترها به روش تابع درستنمایی و $\hat{\theta}$ تخمین غیرمقید پارامتر θ به روش تابع درستنمایی است. ابتدا اشاره به MLE مقید می‌نماییم:

فرض کنیم بطور کلی تابع درستنمایی متغیرتصادفی y را که دارای پارامتر θ می‌باشد را داشته باشیم:

$$H=L(y;\theta)-\lambda' h(\theta)$$

$h(\theta)$ تابعی خطی یا غیرخطی از پارامتر θ می‌باشد که تحت شرایط H_0 داخل داده‌های آماری وجود دارد.

بطوریکه $h(\theta)=0$ بردار G عنصری است که عناصرش توابعی قابل مشتق‌گیری باشند. مثلاً در مثال مدل رگرسیون چند متغیره $h(\theta)=R\beta-r=0$ می‌باشد. λ عبارتست از Shadow Price هر قیدی که در مدل لاگرانژ بکار می‌رود.

شرایط اولیه برای به حداکثر رساندن H عبارتست از:

$$\frac{\partial H}{\partial \theta}=\frac{\partial L(y;\theta)}{\partial \theta}\bigg|_{\hat{\theta}_c}-\frac{\partial h'(\theta)}{\partial \theta}\bigg|_{\hat{\theta}_c}\hat{\lambda}=0$$

$$\frac{\partial H}{\partial \lambda} = h(\hat{\theta}_c) = 0$$

از رابطه اول داریم:

$$\Rightarrow \frac{\partial L(y;\theta)}{\partial \theta}\Big|_{\hat{\theta}_c} = \frac{\partial h'(\theta)}{\partial \theta}\Big|_{\hat{\theta}_c} \hat{\lambda}$$

بنابراین آزمون LM براساس آزمون Rao's Score (1948) به‌صورت زیر می‌باشد:

$$\frac{-\partial L(y;\theta)}{\partial \theta'}\Big|_{\hat{\theta}_c} \left[E\, \frac{-\partial^2 L(\theta)}{\partial\theta \partial\theta'}\Big|_{\hat{\theta}_c} \right]^{-1} \frac{\partial L(y;\theta)}{\partial \theta}\Big|_{\hat{\theta}_c}$$

که تحت شرایط H_0 دارای توزیع مجانبی χ^2 با درجه آزادی G می‌باشد.

برای قید $R\beta = r$ داریم:

$$1) \quad \frac{\partial H}{\partial \beta} = \frac{\partial L(y;\theta)}{\partial \beta} - R'\lambda = \frac{1}{\hat{\sigma}^2} X'(y - X\hat{\beta}_c) - R'\hat{\lambda} = 0$$

$$2) \quad \frac{\partial H}{\partial \sigma^2} = \frac{\partial L(y;\theta)}{\partial \sigma^2} - \frac{N}{2\hat{\sigma}_c^4} + \frac{1}{2\hat{\sigma}_c^4}(y - X\hat{\beta}_c)'(y - X\hat{\beta}_c) = 0$$

$$3) \quad \frac{\partial H}{\partial \lambda} = R\hat{\beta}_c - r = 0$$

تخمین σ^2 به روش تابع درست نمایی عبارتست از:

$$\hat{\sigma}_c^2 = \frac{1}{T}(Y - X\hat{\beta})'(Y - X\hat{\beta}_c) = \frac{1}{T}\hat{U}'\hat{U}$$

اگر رابطه (۱) را در $R(X'X)^{-1}$ پیش ضرب کنیم و آن را برای λ حل کنیم داریم:

$$\hat{\lambda} = \frac{1}{\hat{\sigma}_C^2}\left[R(X'X)^{-1}R'\right]^{-1} R(X'X)^{-1}\frac{\partial L(y;\beta)}{\partial \hat{\beta}_c}$$

$$\hat{\lambda} = \frac{1}{\hat{\sigma}_C^2}\left[R(X'X)^{-1}R'\right]^{-1} R(\hat{\beta} - \hat{\beta}_c)$$

چون:

$$\hat{\beta}_c = \hat{\beta} + (X'X)^{-1}R'\left[R(X'X)^{-1}R'\right]^{-1}(r - R\hat{\beta})$$

اگر بجای آن قراردهیم و ساده کنیم داریم:

مقدار ضریب لاگرانژ عبارتست از:

$$\hat{\lambda} = \frac{1}{\hat{\sigma}_C^2}\left[R(X'X)^{-1}R'\right]^{-1}(R\hat{\beta} - r)$$

اگر $R\beta = r$ باشد شرایط H_0، آنگاه وقتی $n \to \infty$ میل می‌کند.

$$\hat{\sigma}_c^2 \xrightarrow{1.P} \sigma^2$$

$$\Rightarrow \hat{\lambda} \sim N\left[0, \left[\sigma^2 R(X'X)^{-1} R'\right]^{-1}\right]$$

بنابراین بعد ازجایگزینی $\hat{\sigma}_c^2$ بجای σ^2، فرم درجه دوم زیر:

$$LM = \frac{1}{\hat{\sigma}_C^2} (R\hat{\beta} - r) \; \left[R(X'X)^{-1} R'\right]^{-1} (R\hat{\beta} - r) \; \to \chi_G^2$$

آزمون ضریب لاگرانژ است.

فصل دهم

۱-۱۰ واریانس ناهمسانی[۳۷]

شرایطی وجود دارد که واریانس جمله اخلال در مدلهای خطی ثابت نمی‌باشد. برای مثال در تحلیل‌های اقتصادی که از داده‌های مقطع زمانی استفاده می‌شود، واحدهای مورد مطالعه عموماً خانوارها، بنگاهها و افراد می‌باشند و میزانی که مدل خطی می‌تواند رفتار این واحدها را توضیح دهد بستگی به وضعیت و موقعیت خاصی دارد.

مثال ۱:

فرض کنید، برای سال معینی در صنعت مشخص، سودبنگاهها به میزان مخارج تبلیغات و میزان هزینه‌ای که در تحقیقات آنها انجام می‌گردد بستگی دارد.

$$\pi_i = \alpha_1 + \alpha_2 A_i + \alpha_3 R_i + u_i \qquad i = 1, \ldots, n$$

$\pi_i =$ سود بنگاه i ام

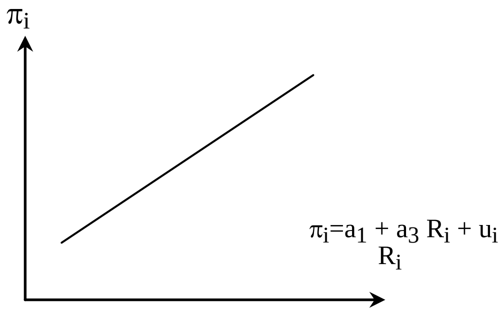
$A_i =$ مخارج تبلیغاتی بنگاه i ام

$R_i =$ مخارج تحقیقات بنگاه i ام

$n =$ تعداد بنگاه‌ها

37 - Seemingly Unrelated Regression (SUR).

هر چند منطقی بنظر می‌رسد که میانگین u_i صفر باشد ولی اندازه واریانس جمله اخلال به بزرگی و کوچکی مقیاس بنگاه مربوط می‌شود که به میزان مخارجی که بنگاهها در بخش تحقیقات بنگاهشان هزینه می‌کنند بستگی دارد. بدین صورت که تغییرات واریانس جمله اخلال برای بنگاههای بزرگ بیشتر از تغییرات واریانس برای بنگاههای کوچک است. چون هزینه نمودن در زمینه تبلیغات و تحقیقات برای موسسات کوچک توام با ریسک و خطر است. لذا ما انتظار داریم که تغییرات اطراف میانگین سود بنگاهها برای موسسات بزرگ بیشتر از تغییرات مشابه برای بنگاههای کوچک باشد. به شکل توجه کنید.

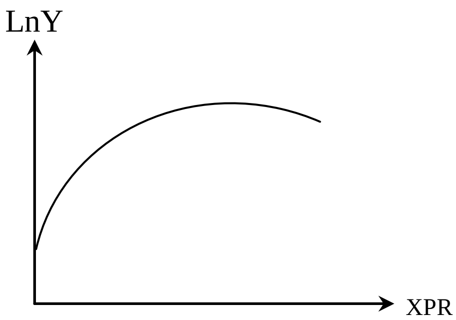


برای میزان معینی مخارج تبلیغات، انتظار داریم که واریانس u_i به R_i مربوط شود. در اینجا ما می‌خواهیم اطلاعات مربوط به بنگاه بزرگ را وزن کمتری از بنگاههای کوچک بدهیم، چون رابطه خطی برای بنگاههای کوچک بطور دقیقتری برقرار است.

مثال ۲:

مینسر Mincer نشان داد که قدرت درآمد افراد تابعی از سالهای تحصیل (S) و تجربه (XPR) در یک شغل است.

$$\text{Ln}\,Y_j = \alpha_\circ + \alpha_1 (XPR)_i - \alpha_2 (XPR)_i^2 + \alpha_3 S_i + \alpha_4 \phi_i + e_i$$



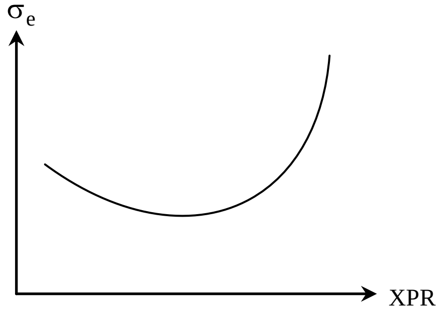
یک ساعت درآمد فرد i ام $\equiv Y_i$

تجربه فرد i ام سالهای تحصیلی- سن $\equiv (XPR)_i$

سالهای تحصیلی فرد i ام $\equiv S_i$

سایر متغیرها مانند نژاد، جنس، عصبانی $= \phi_i$

در اینجا ملاحظه می‌شود که واریانس جمله اخلال تقریباً بصورت فرم درجه دوم تابعی از تجربه است.



$$\sigma_e^2 = a + b(XPR) + c(XPR)^2$$

دلایلی وجود دارد که چرا واریانس جملات اخلال در پایان دوران خدمت بیشتر است.

۱- در مرحله شروع کار حقوق پایین و تحرک بالا در مقابل حقوق بالا و تحرک نسبتاً پایین در سالهای پایانی.

۲- مختلف بودن طرح‌های بازنشستگی

مثال ۳:

خانوارها نیز مانند بنگاهها، اندازه‌های مختلف دارند، بطوریکه اندازه خانوار می‌تواند توسط درآمد یا تعداد اعضای خانواده اندازه‌گیری شود.

فرض کنیم مخارج خانوار توسط مدل زیر توضیح داده شود:

$$E_i = b_1 + b_2 Y_i + b_3 N_i + e_i$$

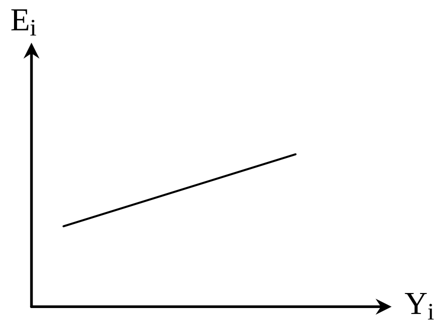
$E_i \equiv$ مخارج خانوار

$Y_i \equiv$ کل درآمد قابل تصرف

$N_i \equiv$ تعداد اعضای خانوار

چون درآمد بالا موجب احتیاط بیشتر خانوارها نسبت به مصرف و پس‌انداز آنهاست: انتظار داریم که وقتی درآمد خانوار بالاست، پراکندگی تغییرات مخارج نیز بیشتر باشد.

یک خانوار نسبتاً ثروتمند، دنبال داشتن منزل بزرگتر با استخر شنا می‌باشد و سعی دارد تعطیلات آخر هفته را به ساحل دریا برود. اگر دیاگرام پراکندگی مشاهدات یک نمونه را برای E_i و Y_i ترسیم کنیم احتمالاً بصورت شکل زیر خواهد شد.



Praise- Houthakou

در مطالعاتشان در سال ۱۹۵۵ در مورد بررسی بودجه خانوار به چنین نتایجی رسیدند.

مثال ۴:

در استفاده از داده‌های آماری غالباً با متوسط داده‌ها مواجه می‌شویم

در یک تابع تولید خطی Y_{it} مقدار محصول بدست آمده در هکتار i ام توسط مزرعه t ام

$$y_{it} = \beta_\circ + \beta_1 X_{1\,it} + \beta_2 X_{2\,it} + u_{it}$$

$X_{1\,it}$: مقدار نیروی کار که توسط مزرعه t ام در هکتار i ام بکار می‌رود

X_{2it} : مقدار نیروی سرمایه که توسط مزرعه t ام در هکتار i ام بکار می‌رود

i.i.d u_{it}

$$E(u_{it})=0 \quad , \quad E(u_{it}^2)=\sigma^2$$

فرض کنید داده‌های آماری ما بصورت $\bar{y}_t$، $\bar{X}_{1t}$ و $\bar{X}_{2t}$ در دست است.

$$\bar{y}_t=\frac{1}{N_t}\sum_{i=1}^{N_t} y_{it}$$

$$\bar{X}_{1t}=\frac{1}{N_t}\sum_{i=1}^{N_t} X_{1it}$$

$$\bar{X}_{2t}=\frac{1}{N_t}\sum_{i=1}^{N_t} X_{2it}$$

N_t تعداد هکتار زمین متعلق به مزرعه t ام

$$\bar{y}_t=\beta_\circ+\beta_1\bar{X}_{1t}+\beta_2\bar{X}_{2t}+\bar{u}_{it} \qquad \bar{U}_t=\frac{1}{N_t}\sum_{i=1}^{N_t} u_{it}$$

$$E(\bar{u}_t)=0$$

$\bar{u}_t$ ها از هم مستقلند ولی برای همه مزرعه‌ها برابر نیستند و مدل دچار واریانس ناهمسانی است.

$$Var(\bar{u}_t)=\frac{\sigma^2}{N_t}$$

مثال ۵:

مدل ضرایب تصادفی[38]

$$y_t=\beta_{\circ t}+\beta_{1t}X_{t1}+\beta_{2t}X_{t2}+\ldots+\beta_{kt}X_{tk}$$

چون ضرایب تصافی هستند نیازی به جمله اخلال نیست.

فرض کنیم: $\beta_{it}=E(\beta_{it})+V_{it} \qquad E(V_{it})=0$

$$\beta_i^*=E(\beta_i^*)+V_{it}$$

$$y_t=\beta_0^*+V_{01}+(\beta_i^*+V_{1t})X_{t1}+\ldots+(\beta_k^*+V_{kt})X_{tk}$$

$$y_t=\sum_{i=0}^{k}\beta_i^* X_{1ti}+\underbrace{\sum_{i=0}^{k} V_{it}X_{it}}_{U_t} \qquad X_{0t}=1$$

$E(u_t)=0$ و چون X_{it} غیرتصادفی

38 - Random Coefficient Model.

فرض کنیم V_{it} ها مستقل از هم باشند و $Var(V_{it})$ ثابت و بستگی به t ندارد.

$$\Rightarrow \quad E(V_{it}^2) = \tau_i^2$$

$$\therefore\ E(u_t^2) = \sum_{i=0}^{k} X_{ti}^2 \tau_i^t$$

$$E(u_t u_s) = E(\sum X_{ti} V_{it} . \sum X_{si} V_{js})$$

$$= E(\sum_{i=j=0}^{K} X_{ti} X_{tj} V_{it} V_{js})$$

$$= \sum_{i=j=0}^{K} X_{ti} X_{tj} E(V_{it} V_{js}) = 0$$

خلاصه آنکه، دلایل زیادی وجود دارد که واریانس جمله های اخلال روی مشاهدات مختلف نبایستی ضرورتاً یکی باشند. هر چند روش OLS بدون تورش خواهد بود ولی کاراً نیست.

بجای درنظر گرفتن تابع زیان[۳۹] $u'u$ باید $u u'\Omega^{-1}$ را در نظر گرفت. بجای آنکه اجازه دهیم هر جمله اخلال دارای وزن برابری در رگرسیون باشد باید وزن آنها را نسبت به تغییراتشان بطور معکوس درنظر بگیریم. هر چقدر u_t تغییراتش بیشتر باشد مشاهدات متناسب با e_t کمتری روی X_t و Y_t در تعیین معادله مدل باید مورد استفاده قرار گیرد.

۱۰-۲ واریانس ناهمسانی گروهی[۴۰]

در خیلی از موارد امکان این وجود دارد که در داخل گروهی از مشاهدات واریانس جمله اخلال ثابت است ولی بین گروه‌ها واریانس متفاوت باشند. مثلاً در بحث هزینه خانوارها ممکن است درآمدهای ۲۰۰۰۰-۰، ۴۹۰۰۰-۲۰۰۰۱ و ۵۰۰۰۱ به بالا، بطوریکه در داخل هر طبقه درآمدی واریانس جمله اخلال ثابت باشد ولی بین گروهها ثابت نباشد.

فرض کنیم T تعداد مشاهدات، باشد و G گروه مشاهده داشته باشیم. به‌نحوی که $T_g > k, T_1 + T_2 + ... + T_G = T$ باشد.

مدل رگرسیون خطی بصورت زیر خواهد شد:

$$y_1 = X_1 \beta + u_1$$

$$y_2 = X_2 \beta + u_2$$

$$\vdots$$

39 -Loss Function.

40 - Grouped Heteroskedasticity.

$$y_g = X_g\,\beta + u_g$$
$$\vdots$$
$$y_G = X_G\,\beta + u_G$$

که u_g, X_g, y_g به ترتیب عبارتند از: مشاهدات روی متغیر وابسته، متغیر مستقل و جمله اخلال در گروه g . توجه کنیم که بردار β برای همه گروه‌ها مشابه هم است.

$$y = X\beta + u$$

$$y' = (y'_1, y'_2, \ldots, y'_G)$$

$$X' = (X'_1, X'_2, \ldots, X'_G)$$

$$u = (u'_1, u'_2, \ldots, u'_G)$$

$$E(u) = 0$$

$$E(uu') = E\begin{bmatrix} u_1 u'_1 & u_1 u'_2 & \cdots & u_1 u'_G \\ \vdots & & & \\ u_G u'_1 & \cdots & & u_G u'_G \end{bmatrix} = \begin{bmatrix} \sigma_1^2 I_{T1} & 0 & 0 \\ & \sigma_g^2 I_{Tg} & 0 \\ 0 & 0 & \sigma_G^2 I_{TG} \end{bmatrix}$$

σ_g^2 واریانس جمله اخلال در داخل گروهg است. اگر مشکل این باشد، GLS متناظر با مدل خواهیم داشت.

$$P' = \begin{bmatrix} \frac{1}{\sigma_1} I_{T1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{\sigma_G} I_{TG} \end{bmatrix}$$

$$P'y = P'X\beta + P'u$$

$$y^* = X^*\beta + e \quad \Rightarrow \quad Ee^*e'^* = I_T$$

بکاربردن OLS در مدل فوق معادل GLS در مدل اصلی است.

تخمین‌زننده β سازگار و بدون تورش، خطی و کارآ است، ولی غالباً σ_g^2 معلوم نیست. لذا اگر $\hat{\beta}_g = (X'_g X_g)^{-1} X'_g y_g$ تخمین OLS گروه g ام باشد.

$$\sigma_g^2 = \frac{(y_g - X_g\hat{\beta}_g)'(y_g - X_g\hat{\beta}_g)}{T_g - K}$$

آنگاه:
$$\hat{P}'=\begin{bmatrix}\frac{1}{\hat{\sigma}_1}I_{T1} & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & \frac{1}{\hat{\sigma}_G}I_{TG}\end{bmatrix}$$

سپس $\tilde{\tilde{\beta}}=\left[(\hat{P}'X)'(\hat{P}'X)\right]^{-1}(\hat{P}'X)'\hat{P}'y$ یک Feasible GLS است برای زمانیکه، واریانس ناهمسانی بین گروهها داریم.

۱۰-۳ تبعات واریانس ناهمسانی درروش OLS برروی تخمین زننده‌ها

چنانچه واریانس ناهمسانی بصورت زیر باشد:

$$y_i=\beta x_i+\varepsilon$$

$$Var(\varepsilon_i)=\sigma^2 x_i^2$$

$$OLS: Var(\hat{\beta}_{OLS})=\sigma^2[x'x]^{-1}[x'\Omega x]\;[x'x]^{-1}=\sigma^2\frac{\sum x_i^4}{\left(\sum x_i^2\right)^2}$$

$$GLS: Var(\hat{\beta}_{GLS})=\sigma^2[x'\Omega^{-1}x]^{-1}=\frac{\sigma^2}{n}$$

$$K=\frac{\frac{\Sigma x_i^4}{(\Sigma x_i^2)^2}\sigma^2}{\frac{\sigma^2}{n}}=\frac{n\Sigma\chi_i^4}{(\Sigma x_i^2)^2}=\frac{\left[\frac{\Sigma x_i^4}{n}\right]}{\left[\frac{\Sigma x_i^2}{n}\right]^2}$$

$$x_i^2=a_i$$

$$K=\frac{\frac{1}{n}\sum a_i^2}{\left[\frac{1}{n}\sum a_i\right]^2}=\frac{\frac{1}{n}\sum a_i^2}{\bar{a}^2}=\frac{\frac{1}{n}\sum(a_i-\bar{a})^2+\bar{a}^2}{\bar{a}^2}=1+\frac{\frac{1}{n}\sum(a_i-\bar{a})^2}{\bar{a}^2}>1$$

OLS هرگز تخمین زننده کارایی همانند GLS نخواهد بود.

اگر x_i دارای میانگین صفر وگشتاور چهارم معینی باشد، آنگاه:

$$P\lim K=\frac{Ex_i^4}{E\left(x_i^2\right)^2}=\frac{\mu_4}{\sigma_x^4}$$

۱۰-۴ انواع واریانس ناهمسانی

واریانس ناهمسانی بصورت تابعی از متغیرهای مستقل یا برونزا

الف) در معادله $\sigma_i^2=a+b(XPR)_i+C(XPR)_i^2$ یعنی واریانس تابعی خطی از متغیرهای مستقل است.

لذا بطور کلی ممکن است $\sigma_i^2 = Z_i\,\alpha$ باشد. که Z_i میتواند مجموعه‌ای از متغیرهای مستقل و برونزا باشد

$$Z_i = (1, XPR_i, (XPR_i)^2) \qquad \alpha = (a, b, c)$$

ب) Praise – Houthakou (1955) در داده‌های مصرف جاری نشان دادند که واریانس تابعی از مجذور میانگین تابع رگرسیون است.

$$\sigma_i^2 = \sigma^2 (Ey_i)^2$$

$$= \sigma^2 (\beta_1 x_{i1} + \ldots + \beta_k x_{ik})^2$$

یا:

$$\sigma_i = \sigma\beta_1 x_{i1} + \ldots + \sigma\beta_k x_{ik}$$

$$\sigma_i = \beta_1^* x_{i1} + \ldots + \beta_k^* x_{ik}$$

در اینجا انحراف معیار تابعی است از متغیرهای برونزا

پس:

$$Z_i = (1, \ldots, x_{ik}) \qquad \alpha = (\beta_1^*, \ldots, \beta_k^*)$$

ج) در تابع سود واریانس کاملاً مربوط به مخارج تحقیقات می‌شود این شکل ممکن است بصورت $\sigma_i^2 = \sigma^2 (R_i)^\lambda$ باشد. در اینجا واریانس جمله اخلال بصورت فرم حاصلضربی در پارامتر λ است.

پس فرم سوم:

$$\sigma_i^2 = e^{Z_i\alpha}$$

$$Z_i = (1, LnR_i) \qquad \alpha' = (Ln\sigma^2, \lambda)'$$

حال متناظر با هر یک از حالات پیش آمده، GLS مربوط را مورد استفاده قرار می‌دهیم.

۱- چنانچه وضعیت داده‌های آماری با حالت الف منطبق باشد، $\tilde{\beta}$ را بصورت زیر تعریف می‌کنیم:

$$\tilde{\beta} = (X' D_1^{-1} X)^{-1} X' D_1^{-1} y$$

$$= \left(\sum_{i=1}^{T} (Z_i \hat{\alpha})^{-1} X_i' X_i \right)^{-1} \sum_{i=1}^{T} (Z_i \hat{\alpha})^{-1} X_i' y_i$$

D1 ماتریس قطری است که روی قطر آن $Z_i\alpha$ قرار گرفته است. اما مشکل اینجاست که در غالب اوقات α ها معلوم نیستند و باید تخمین‌زده شوند. یک روش آنست که با استفاده از $\hat{u}_i^2$، α راتخمین بزنیم:

$$\hat{u}_i^2 = Z_i\alpha + v_i \qquad v_i = \hat{u}_i^2 - \sigma_i^2$$

نشان داده خواهد شدکه $Ev_i \neq 0$ و دارای مشکل واریانس ناهمسانی وهمبستگی پیاپی هم خواهیم شد. درنتیجه $\hat{\alpha}_{OLS}$ تورش خواهد داشت. اما اگر حجم مشاهدات باندازه کافی بزرگ باشد α یک تخمین زننده سازگار است ولذا می‌توان به Feasible GLS دست یافت.

Goldfeld- Quandt(1972) تخمین α رابصورت زیر پیشنهاد دادند:

$$\hat{\alpha} = (\sum_{i=1}^{t} Z_i' Z_i)^{-1} \sum_{i=1}^{T} Z_i \hat{e}_i^2 = (Z'Z)^{-1} Z' \hat{e}^2$$

Feasible GLS:

$$\tilde{\tilde{\beta}} = (X' D_1^{-1} X)^{-1} X' D_1^{-1} y$$

$$= \left(\sum_{i=1}^{T} (Z_i \alpha)^{-1} X_i' X_i \right)^{-1} \sum_{i=1}^{T} (Z_i \alpha)^{-1} X_i' y_i$$

این تخمین زننده سازگار ، وبطور مجانبی دارای توزیع نرمال است واگر u_i دارای توزیع نرمال باشد بطور مجانبی $\tilde{\tilde{\beta}}$ کارآ هم خواهدبود.

Amemiya (1977) پیشنهاد نمود:

الف- $\hat{\alpha}$ را با روش OLS تخمین بزنید.

ب- $\hat{\Sigma}_1$ را تشکیل دهید بنحوی که یک ماتریس قطری باشد که عناصر روی قطر عبارت باشداز:

چون: $EV_i^2 = 2\sigma^4$ $\quad 2\hat{\alpha}_i^4 = 2(Z_i\hat{\alpha})^2$

ج- $\hat{\hat{\alpha}} = (Z' \hat{\sum}_1^{-1} Z)^{-1} Z' \hat{\sum}_1^{-1} \hat{u}^2$

$$= \left(\sum_{i=1}^{T} (Z_i\hat{\alpha})^{-2} Z_i' Z_i \right)^{-1} \sum_{i=1}^{T} (Z_i\hat{\alpha})^{-2} Z_i' \hat{u}_i^2$$

این تخمین‌زننده دارای ماتریس واریانس کواریانس $(Z' \sum_1^{-1} Z)^{-1}$ است که حد پایین R.C را تامین می‌کند. لذا وقتی واریانس ناهمسانی بصورت $\sigma_i^2 = Z_i\alpha$ باشد:

$$\tilde{\tilde{\beta}} = (X' \hat{\hat{D}}_1^{-1} X)^{-1} X' \hat{\hat{D}}_1^{-1} y$$

$$= (\sum_{i=1}^{t} (Z_i\hat{\hat{\alpha}})^{-1} X_i' X_i)^{-1} \sum_{i=1}^{t} (Z_i\hat{\hat{\alpha}})^{-1} X_i' y_i$$

Feasible_GLS است.

۲- انحراف ازمعیار تابعی خطی ازمتغیرهای برونزا است.

$$\sigma_i = Z_i\alpha$$

$$y_i = X_i\beta + e_i$$

$$E(u_i) = 0, E({u_i}^2) = \sigma_i^2 = (Z_i\alpha)^2, E(u_i u_j) = 0 \quad i \neq j$$

$$u_i \sim N(0, \sigma_i^2)$$

$$\tilde{\tilde{\beta}} = (X' D_2^{-1} X)^{-1} X' D_2^{-1} y$$

$$= \left[\sum_{i=1}^{T} (Z_i\alpha)^2 X_i' X_i \right]^{-1} \sum_{i=1}^{T} (Z_i\alpha)^{-2} X_i' y_i$$

D_2 ماتریسی قطری است که عناصر روی قطرش عبارتست از: $(Z_i\alpha)^2$.

ملاحظه می‌شود که بحث تخمین FGLS مدل قبلی قابل اجراست، با این تفاوت که تخمین دیگری برای α بدست خواهد آمد.

فرض کنید $\frac{u_1}{\sigma_1},\frac{u_2}{\sigma_2},....,\frac{u_T}{\sigma_T}$ ها بطور مستقل ویکنواخت با میانگین ۰ و واریانس ۱ توزیع شده باشند، میانگین قدر مطلق این متغیر استاندارد نرمال یعنی: $\frac{u_i}{\sigma_i}$ عبارتست از:

$$E\frac{|u_i|}{\sigma_i}=2\int_{\circ}^{\infty}\omega f(\omega)d\omega\equiv c$$

$f(\omega)$ عبارت است ازتابع احتمال چگالی $\omega=\frac{u_i}{\sigma_i}$ و c ثابت است.

فرض شده که $f(\omega)$ نسبت به صفر متقارن است.

بطور مثال اگر $u_i\sim N(0,\sigma_i^2)\Rightarrow c=(\frac{2}{\pi})^{1/2}$

بنابراین: $E|u_i|=c\sigma_i=cZ_i\alpha$

$$|u_i|=cZ_i\alpha+v_i$$

Gleisjer(1969) پیشنهاد کرد $C\alpha$ را بکار ببریم و تخمین آن خواهد شد:

$$(\hat{C\alpha})=(\sum_{i=1}^{T}Z_i'Z_i)^{-1}\sum_{i=1}^{T}Z_i'|\hat{u}_i|=(Z'Z)^{-1}Z'|\hat{u}|$$

$$\tilde{\tilde{\beta}}_{FGLS}=(\sum(Z_i\hat{C\alpha})^{-2}X_i'X_i)^{-1}\sum_{i=1}^{T}(Z_i\hat{C\alpha})^{-2}X_i'Y_i$$

توجه کنیم که در نمونه‌های کوچک $\hat{C\alpha}$ یک تخمین زننده با تورش و غیرکاراً است چون نشان داده می‌شود V_i دارای میانگین صفر نیست و دارای واریانس ناهمسانی و هم بستگی پیاپی است. روش جایگزین و کاراتر FGLS است. مراحل کار بصورت زیر است:

الف- تخمین $C\alpha$ رابدست می‌آوریم اگر c معلوم باشد $\hat{\alpha}=\frac{c\hat{\alpha}}{c}$

ب- $\hat{\sum}_2$ را تشکیل می‌دهیم، یک ماتریس واریانس کواریانس است که قطری است و اجزای روی قطر آن عبارتست از واریانس σ_i^2 که عبارت از: $(1-c^2)(Z_i\hat{\alpha})^2=EV_i^2$

ج- FGLS ، $C\alpha$ را تشکیل می‌دهیم:

$$\hat{\hat{C\alpha}}=(Z'\sum_2^{-1}Z)^{-1}Z'\hat{\sum}_2^{-1}|\hat{u}|$$

در نهایت $\tilde{\beta}_{FGLS}$ خواهد شد.

$$\tilde{\tilde{\beta}}=(\sum_{i=1}^{T}(Z_ic\hat{\hat{\alpha}})^{-2}X_i'X_i)^{-1}\sum_{i=1}^{T}(Z_ic\hat{\hat{\alpha}})^{-2}X_i'Y_i$$

نکته قابل توجه اینست که نیاز به c داریم لذا محقق باید بتواند توزیع u_i را مشخص کند.

Harvey (1974) نشانداد که $c\hat{\alpha}$ و $c\hat{\hat{\alpha}}$ هر دو سازگار هستند و دارای ماتریس واریانس کواریانس‌های زیر:

$$V_{c\hat{\alpha}} = (Z'Z)^{-1}Z^{-1}\sum\nolimits_2 Z(Z'Z)^{-1}$$

$$V_{C\hat{\alpha}} = (Z'\textstyle\sum_2^{-1} Z)^{-1}$$

عناصر روی $\Sigma_2 = (1-c)(Z_i\alpha)^2$

اولین واریانس یک تخمین زننده OLS است، دوّمی تخمین زننده GLS است. اگر u_i دارای توزیع نرمال باشد دوّمی حد پائین R.C را تامین می‌کند.

پس اگر توزیع u_i نرمال باشد هر دو $\tilde{\tilde{\beta}}$ تخمین زننده بطور مجانبی کارآ هستند و در نمونه‌های کوچکتر بهتر است دوّمی مورد استفاده قرار گیرد.

$$Vc(\tilde{\tilde{\beta}}) = \frac{1}{c^2}\left[\sum_{i=1}^{T}(Z_i\hat{\hat{C}})^{-2}X_i'X_i\right]^{-1} = \left[\sum_{i=1}^{T}(Z_i\hat{\hat{\alpha}})^{-2}X_i'X_i\right]^{-1}$$

۳- Harvey (1976) فرم کلی‌تری را برای واریانس ناهمسانی بررسی کرد و آن واریانس ناهمسانی به صورت حاصلضربی است.

$$E(u_i^2) = \sigma_i^2 = e^{Z_i\alpha} \qquad u_i \sim N$$

یعنی:

$$\sigma_i^2 = e^{Z_{i1}\alpha_1}.\,e^{Z_i\alpha_2}\ldots e^{Z_{i1}\alpha_3}$$

بطور مثال:

$$\alpha = (\mathrm{Ln}\sigma^2, \lambda_1, \lambda_2)'$$

$$Z_i = (1, \mathrm{Ln}x_{i1}, \mathrm{Ln}x_{i2})$$

$$\Rightarrow \sigma_i^2 = \sigma^2 x_{i1}^{\lambda_1},\ \sigma^2 x_{i2}^{\lambda_2}$$

اگر D_3 ماتریس قطری باشد که عناصر روی قطر آن $e^{Z_{i1}}$ تخمین زننده GLS خواهد شد.

$$\tilde{\beta} = (X'D_3^{-1}X)^{-1}X'D_3^{-1}y$$

$$= \left[\sum^{T}[e^{Z_i\alpha}]^{-1}X_i'X_i\right]^{-1}\left[\sum_{i=1}^{T}[e^{Z_i\alpha}]^{-1}X_i'Y_i\right]^{-1}$$

که میانگین آن β و واریانس آن برابر است با:

$$(X'D_3^{-1}X)^{-1}$$

اما مجدداً مشکل مجهول بودن α مطرح است. فرم فوق را بصورت زیر می‌نویسیم:

$$\mathrm{Ln}\sigma_i^2 = Z_i\alpha$$

$$\mathrm{Ln}\,\hat{u}_i^2 = Z_i\alpha + v_i \qquad ; \qquad v_i = \mathrm{Ln}\left(\frac{\hat{u}_i^2}{\sigma_i^2}\right)$$

$$\hat{\alpha} = \left[\sum_{i=1}^{T} Z_i' Z_i\right]^{-1} \sum_{i=1}^{T} Z_i' \mathrm{Ln}\,\hat{u}_i^2$$

$$\tilde{\tilde{\beta}} = \left[\sum_{i=1}^{T} \left[\exp(Z_i\hat{\alpha})\right]^{-1} X_i' X_i\right]^{-1} \sum_{i=1}^{T} \left[\exp(Z_i\hat{\alpha})\right]^{-1} X_i' Y_i$$

این تخمین زننده سازگار بوده و بطور مجانبی نرمال توزیع شده و بطور مجانبی کارآ است.

۱۰-۵ -۱ آزمون تشخیص وجود واریانس ناهمسانی

۱- آزمون برای تشخیص وجود واریانس ناهمسانی گروهی نسبت ماکزیمم لایکلی هود:

$$H_0: \sigma_1^2 = \sigma_2^2 = \ldots = \sigma^2$$
$$H_1: \text{not all } \sigma_1^2 = \sigma^2$$

فرض کنیم u_i ها بطور نرمال توزیع شده باشند. می‌توان ثابت نمود که بطور مجانبی

$$-2\mathrm{Ln}\frac{L(\hat{\omega})}{L(\hat{\Omega})} \sim \chi_q^2$$

Ω نشان دهنده K+G متغیر β_k و σ_g هاست .

ω نشانگر K+G-(G-1)، β_kها و σ^2 است که تحت شرایط H_0 بدست می‌آیند.

$$L(\omega) = L(y\,|\,X,\beta;\ \sigma_1^2 = \sigma_2^2 = \ldots \sigma^2)$$

$$= \frac{1}{(2\pi)^{T/2}\sigma^T} \exp\left\{-\frac{1}{2\sigma^2}(Y - X\beta)'(Y - X\beta)\right\}$$

$$L(\omega) = \frac{1}{(2\pi)^{T/2}\hat{\sigma}^T} \exp\left\{-\frac{1}{2\hat{\sigma}^2}(Y - X\hat{\beta})'(Y - X\hat{\beta})\right\}$$

$$\hat{\beta} = (X'X)^{-1}X^{-1}Y$$

$$\hat{\sigma}^2 = \frac{(Y - X\hat{\beta})'(Y - X\beta)}{T}$$

$$L(\Omega) = L(y\,|\,X,\beta,\ \sigma_1^2, \sigma_2^2 \ldots \sigma_G^2)$$

$$= \frac{1}{(2\pi)^{T/2}\sigma_1^{T_1}\sigma_2^{T_2}\ldots\sigma_G^{T_G}} \exp\left\{-\frac{1}{2}\sum_{g=1}^{G}\frac{1}{\sigma_g^2}(Y_g - X_g\beta)'(Y_g - X_g\beta)\right\}$$

$$L(\hat{\Omega})=\frac{1}{(2\pi)^{T/2}\ \tilde{\sigma}_1^{T_1}\ldots\tilde{\sigma}_G^{T_G}}\exp\left\{-\frac{1}{2}\sum_{g=1}^{G}\ \frac{1}{\tilde{\sigma}_g^2}(Y_g-X_g\tilde{\tilde{\beta}})'(Y_g-X_g\tilde{\tilde{\beta}})\right\}$$

$$\tilde{\tilde{\beta}}=(X'SX)^{-1}X'^{-1}S^{-1}Y$$

بطوریکه:

$$S=\begin{bmatrix}\tilde{\sigma}_1 I_{T_1} & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & \tilde{\sigma}^2{}_G\ I_{TG}\end{bmatrix}$$

$$\tilde{\sigma}_g^2=\frac{(y_g-X_g\tilde{\tilde{\beta}}_g)'(y_g-X_g\tilde{\tilde{\beta}}_g)}{T_g}\qquad g=1,2,\ldots,G$$

$\tilde{\tilde{\beta}}_g=(X_g'X_g)^{-1}X_g'Y_g$ که بعنوان مقادیر اولیه برای $\tilde{\sigma}_g^2$ بکار برده می‌شود.

این روش تکرار می‌شود تا تفاوت بین آنها کم بشود.

$$l=\frac{L(\hat{\omega})}{L(\hat{\Omega})}=\frac{\tilde{\sigma}_1^{T_1}\ \ \tilde{\sigma}_2^{T_2}\ldots\tilde{\sigma}_G^{T_G}}{\tilde{\sigma}^{T}}$$

$$-2Lnl=-2(T_1Ln\tilde{\sigma}_1+T_2Ln\ \tilde{\sigma}_2+\ldots+T_GLn\tilde{\sigma}_G)+2TLn\tilde{\sigma}$$

$$=2\left[TLn\tilde{\sigma}-(T_1Ln\tilde{\sigma}_1+T_2Ln\ \tilde{\sigma}_2+\ldots+T_GLn\tilde{\sigma}_G)\right]\sim\chi^2_{G-1}$$

اگر بتوانیم بطور مجانبی به $\tilde{\sigma}_i$ ها دسترسی پیدا کنیم، می‌توانیم از $\tilde{\sigma}_i$ هایی که سازگارند ولی غیرکارآ هستند با استفاده از $\hat{\beta}_g=(X_g'X_g)^{-1}X_g'Y_g$ در $\tilde{\sigma}_g^2$ استفاده کنیم. آماره‌ای که از این طریق محاسبه می‌شود دارای توزیع χ^2_{G-1} است ولی قدرتش در نمونه‌های کوچک کمتر از زمانی است که از $\tilde{\sigma}$ استفاده می‌کنیم.

۱۰-۵-۲ آزمون گلدفلد - کوانت[41]

اگر فرضیه H_0 و H_1 را به صورت زیر تعریف کنیم:

H_0 : واریانس همسانی

H_1: واریانس ناهمسانی

آنگاه برای تشخیص وجود واریانس همسانی:

الف) مشاهدات را طبق مقادیر Z مرتب می‌کنیم. Z متغیری است که جمله اخلال بطور فزاینده با آن در ارتباط است.

41 - Goldfield - Quandt.

ب) P تا مشاهده را از مرکز انتخاب و حذف می‌کنیم.

ج) دو تا رگرسیون با تعداد $\frac{T-P}{2} > K$ انجام می‌دهیم.

د) RSS_1 گروه کوچک و RSS_2 گروه بزرگ را محاسبه می‌کنیم.

یعنی اگر با Z افزایش پیدا می‌کند، RSS_1 مجموع مجذورات جملات پسماند مربوط به گروه مشاهداتی است که دارای مقادیر کوچک Z می‌باشد. نسبت S به صورت زیر دارای توزیع F میباشد.

$$S = \frac{RSS_2}{RSS_1} \sim F_{\frac{T-P-2k}{2}, \frac{T-P-2k}{2}}$$

تحت شرایط H_0 که واریانس همسانی است. H_0 رد می‌شود یعنی اگر $S>F_\alpha$، مدل دچار واریانس ناهمسانی است.

اگر در مدل واریانس ناهمسانی وجود داشته باشد، F بزرگ می‌شود. اگر F محاسبه شده کوچکتر از F جدول باشد آنوقت واریانس همسانی پذیرفته می‌شود، در غیر اینصورت واریانس ناهمسانی پذیرفته می‌شود.

۱۰-۵-۲-۱ اشکالات وارد به روش آزمون فوق

۱- قدرت آزمون بستگی به عدد P دارد اگر P بزرگ باشد، قدرت آزمون کاهش پیدا می‌کند چون تعداد مشاهدات در هر قسمت کاهش پیدا می‌کند، ولی همزمان اگر P را کوچک انتخاب کنیم وجود مشاهدات مرکزی باعث می‌شود که دو RSS خیلی از یکدیگر متفاوت نباشند.

۲- قدرت آزمون همچنین به پراکندگی Z_i نسبت به میانگین دارد. هر چند انحراف از میانگین Z بزرگتر باشد قدرت آزمون بالاتر است.

۳- اگر P=0 باشد این آزمون به طور مجانبی تبدیل به آزمون ML (واریانس ناهمسانی گروهی) خواهد شد که فرض کرده‌ایم دو گروه داریم (G=2).

۴- چون u را نرمال فرض می‌کنیم این نسبت خوبی است برای مشاهدات محدود اگر فرض نرمال بودن را لغو کنیم باید دنبال نتایج بطور مجانبی باشیم.

۱۰-۵-۳ آزمون بریوش - پایگان[۴۲]: (۱۹۷۹)

آنها واریانس ناهمسانی را بصورت وسیعتری در نظر گرفتند. یعنی:

$$\sigma_i^2 = h(Z_i \alpha)$$

حالتهای قبلی شکلی خاص از این تابع هستند و واریانس ناهمسانی گروهی نیز حالتی است که Z شامل متغیرهای مجازی مخصوص است.

آنها تابع تخمین زیر را در نظر گرفتند:

$$\frac{\hat{u}_i^2}{\bar{\sigma}^2} = Z_i \alpha + v_i \qquad (۱)$$

که $\hat{u}$ جمله پسماند، از روش OLS است و $\bar{\sigma}^2 = \sum_{i=1}^{T} \hat{u}_i^2 / T$

42 - Breusch – Pagan.

$\hat{\alpha}$ که از رابطه (۱) بدست می‌آید سازگار است، مگر در مواردی که $h(.)$ معلوم نیست، FGLS را نمی‌توان مورد استفاده قرار داد.

چنانچه u_i بصورت نرمال توزیع شود در این مدل چون اولین جمله Zi ، یک است آزمون به صورت روبرو است: $\alpha=(\alpha_2,\alpha_3,...,\alpha_p)'=0$

فرض کنید ESS (مجموع تغییرات توضیح داده شده) بدست آمده از روش OLS باشد.

$$y_i=\frac{\hat{u}_i^2}{\hat{\sigma}^2} \quad \text{و} \quad \bar{y}=\sum_{i=1}^{T} y_i/T$$

$$\hat{y}=Z_1\alpha \quad \hat{y}=Z_1\hat{\alpha} \quad \Rightarrow \quad \sum_{i=1}^{T}(\hat{y}_i-\bar{y}_i)^2$$

B.P نشان دادند که اگر $(\alpha_2,\alpha_3,...,\alpha_p)=0$ باشد، آنگاه:

$$\frac{1}{2}ESS \overset{asy}{\sim} \chi^2_{p-1}$$

بنابراین، اگر نااطمینانی وجود دارد که کدام یک از فرم‌های گفته شده صحیح است و یا اینکه مایل هستیم $\sigma_i^2=h(Z_i\,\alpha)$ باشد، این آزمون برای واریانس همسانی وجود دارد.

پس اگر مدل مقابل را در نظر بگیریم:

$$Y=X\beta+u \quad Eu=0\ ,\ Eu_t^2=h(Z_i'\alpha)$$

فرضیه H_0 بصورت زیر است:

$$H_0:\alpha_2=\alpha_3=...,=0$$

روش کار بصورت زیر خواهد بود.

(۱) رگرس Y روی X_t و $\hat{u}_{OLS}$ و $\hat{\beta}_{OLS}$ را بدست می‌آوریم.

(۲) $\hat{\sigma}^2=\frac{1}{T}\sum_{t=1}^{T}\hat{u}_t^2$ محاسبه می‌شود.

(۳) $\frac{\hat{u}_t^2}{\hat{\sigma}^2}=y_t$ محاسبه می‌شود.

(۴) $y_t's$ روی Z_t' ها به روش OLS رگرس می‌نماییم آنگاه:

$$\frac{1}{2}ESS=\frac{1}{2}(g'Z(Z'Z)^{-1}Z'g-\frac{1}{T}(i'g)^2\sim\chi^2_{p-1}$$

۱۰-۵-۴ آزمون (۱۹۶۹) پارک - گلسجر[43]

مدل $Y=X\beta+U$ را در نظر می‌گیریم، بطوریکه: $E(u)=0$

$$E(uu')=\sigma^2\,diag(w_1^2,w_2^2,...,w_T^2)$$
$$w_t^2=Eu_t^2=\alpha_1+\alpha_2 Z_t+\alpha_3 Z_t^2$$

این معادله را می‌توانیم از معادلات کلی‌تری بدست آوریم:

$$|u_t|=\gamma+\delta Z_t+V_t$$

---

43 - Park - Glesjer.

ثابت $E(V_t)=0 \quad , \quad E(V_t^2)=$

$$\left|u_t\right|^2=\gamma^2+\delta_t^2+V_t^2+2\gamma\delta Z_t+2(\gamma+\delta Z_t)V_t$$

$$E\left|u_t\right|^2=\gamma^2+\delta^2 Z_t^2+\sigma_v^2+2\gamma\delta Z_t+2(\gamma+\delta Z_t)EV_t$$

$$=(\gamma^2+\sigma_v^2)+2\gamma\delta Z_t+\delta^2 Z_t^2$$

$H_0:\delta=0$ واریانس همسانی

$H_1:\delta\neq 0$ واریانس ناهمسانی

Glesjer مراحل کار را بصورت زیر شرح می‌دهد:

۱- $|\hat{u}_t|$ را روی جمله ثابت و Z_t رگرس می‌کنیم و δ و γ را تخمین می‌زنیم.

۲- نسبت t را بصورت $\frac{\hat{\delta}}{Var(\hat{\delta})}$ تشکیل می‌دهیم.

چنانچه $|t|\geq C_\alpha=(1-\frac{\alpha}{2})$ باشد H_0 رد می‌شود. می‌توان نشان داد که $\hat{\delta}$ بطور مجانبی نرمال است و $t\rightarrow N(0,1)$ می‌باشد.

برخی توصیه می‌کنند که $|\hat{u}_t|$ را برروی X_{t-1} ، X_{t1}^1 ، X_{t1}^2 نیز رگرس کنیم.

۱۰-۵-۵ آزمون وایت[۴۴]

در این آزمون فرضیه H_0 واریانس همسانی است.

به ازای همه i ها $H_0:\sigma_i^2=\sigma^2$

$H_1:$ Not H_0

ابتدا مجذور جمله پسماند مدل رگرسیون را برروی توان اول، دوم و حاصلضرب متغیرهای مستقل مدل رگرس می‌کنیم. R^2 این رگرسیون را بدست می‌آوریم. نشان داده شده است TR^2 دارای توزیع مجانبی χ^2 است. درجه آزادی تعداد متغیرهای مستقل بکار برده شده در این مدل با در نظر گرفتن عرض از مبدا، منهای یک می‌باشد.

این آزمون کلی است و هیچ حالت خاصی از واریانس ناهمسانی را بررسی نمی‌کند. در صورتیکه TR^2 بزرگتر از χ^2 جدول باشد، مسئله واریانس همسانی منتفی است.

۱۰-۵-۶ آزمون ضریب لاگرانژ برای واریانس ناهمسانی گروهی برای ۵ بنگاه با T مشاهده

$$Ln\,L=-\frac{n}{2}TLn\,2\pi-\frac{T}{2}\Sigma\,Ln\sigma_i^2-\frac{1}{2}\,\Sigma\,\frac{\varepsilon_i'\varepsilon_i}{\sigma_i^2}$$

44 - White.

$$H_0: \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_5^2 = \sigma^2$$

بدون اعمال قید یعنی واریانسهای برابر داریم:

$$\frac{\partial \ln L}{\partial \beta} = \Sigma \frac{1}{\sigma_i^2} X_i' (y_i - X_i \beta) = \Sigma \frac{1}{\sigma_i^2} X_i' \varepsilon_i = 0 \qquad i = 1,\ldots,5 \qquad (۱)$$

$$\frac{\partial \ln L}{\partial \sigma_i^2} = -\frac{T}{2\sigma_i^2} + \frac{1}{2\sigma_i^4} \varepsilon_i' \varepsilon_i = 0 \qquad (۲)$$

$$\frac{\partial^2 \ln L}{\partial \beta \partial \beta'} = -\Sigma \frac{1}{\sigma_i^2} X_i' X_i \qquad (۳)$$

$$\frac{\partial^2 \ln L}{\partial (\sigma_i^2)^2} = -\frac{T}{2\sigma_i^4} - \frac{\varepsilon_i' \varepsilon_i}{\sigma_i^6} \qquad (۴)$$

$$\frac{\partial^2 \ln L}{\partial \beta \partial \sigma_i^2} = -\Sigma \frac{1}{\sigma_i^4} X_i' \varepsilon_i \qquad (۵)$$

$$LM = g' V^{-1} g$$

بطوریکه g بردار مشتقات مرتبه اول و V^{-1} حد پایین رائو کرامر تحت شرایط H_0 می‌باشند.

$$g = \begin{bmatrix} \frac{1}{\sigma^2} \Sigma X_i' \varepsilon_i \\ -\frac{T}{2\sigma^2} + \frac{\varepsilon_i' \varepsilon_i}{2\sigma^4} \end{bmatrix}_i$$

$$V = \begin{bmatrix} \frac{1}{\sigma^2} X'X & 0 \\ 0 & \frac{T}{2\sigma^4} \end{bmatrix}$$

معادله دوم را می‌توانیم بصورت زیر بنویسیم:

$$\frac{\partial \ln L}{\partial \sigma_i^2} = \frac{T}{2\sigma^2} \left[\frac{\hat{\sigma}_i^2}{\hat{\sigma}^2} - 1 \right]$$

از طرفی تخمین واریانس مشترک σ^2 از ML عبارتست از:

$$\hat{\sigma}^2 = \frac{\hat{\varepsilon}' \hat{\varepsilon}}{NT} = \frac{\Sigma \hat{\varepsilon}_i' \hat{\varepsilon}_i}{NT}$$

$$\hat{\sigma}^2 = \frac{1}{N} \Sigma \hat{\sigma}_i^2$$

یعنی متوسط ۵ تخمین‌زننده سازگار

$$LM = g' V^{-1} g = \left[\frac{1}{\sigma^2} \Sigma X_i' \varepsilon_i \quad \frac{T}{2\hat{\sigma}^2} + \left(\frac{\hat{\sigma}_i^2}{\hat{\sigma}^2} - 1 \right) \right] \begin{bmatrix} \hat{\sigma}^2 (X'X)^{-1} & 0 \\ 0 & \frac{2\hat{\sigma}^4}{T} \end{bmatrix} \begin{bmatrix} \frac{1}{\hat{\sigma}^2} \Sigma X_i' \varepsilon_i \\ \frac{T}{2\hat{\sigma}^2} \left(\frac{\hat{\sigma}_i^2}{\hat{\sigma}^2} - 1 \right) \end{bmatrix}$$

$$= \left[(X'X)^{-1} \Sigma\ X_i' \varepsilon_i \quad \hat{\sigma}^2 \left(\frac{\hat{\sigma}_i^2}{\hat{\sigma}^2} - 1 \right) \right] \begin{bmatrix} \frac{1}{\hat{\sigma}^2} \Sigma\ X_i' \varepsilon_i \\ \frac{T}{2\hat{\sigma}_i^2} \left(\frac{\hat{\sigma}_i^2}{\sigma^2} - 1 \right) \end{bmatrix}$$

$$\mathrm{LM} = \left[\frac{1}{\hat{\sigma}^2} (X'X)^{-1} \left(\Sigma\ X_i' \varepsilon_i \right)^2 + \frac{T}{2} = \Sigma \left(\frac{\hat{\sigma}_i^2}{\hat{\sigma}^2} - 1 \right) \right] \sim \chi_n^2$$

پس از جایگذاری از ننتایج بدست آمده داریم:

$$= 41.549 > \chi_5^2 = 15.056$$

در این جا فرض نرمال بودن ε_i وجود دارد و بطور جانبی این آزمون با LR برای فرضیه H_0 نیز رد می‌شود.

مثال - آزمون نسبت لاکلیهود برای مسئله فوق

$$LR = -2Ln \frac{L(\hat{w})}{L(\hat{\Omega})} = -2Lnl$$

$$= n T \ln \hat{\sigma}^2 - \Sigma T \ln \hat{\sigma}_i^2 \sim \chi_{n-1}^2$$

$$\hat{\sigma}^2 = \frac{\hat{\varepsilon}_i' \hat{\varepsilon}_i}{NT} \quad , \quad \hat{\sigma}_i^2 = \frac{\hat{\varepsilon}_i' \hat{\varepsilon}_i}{T}$$

چنانچه از OLS استفاده شود

$$-2Lnl = 100 \ln \hat{\sigma}^2 - 20 \Sigma \ln \hat{\sigma}_i^2 = 108.5$$

چنانچه از ML استفاده شود:

$$-2Lnl = 100 \ln \hat{\sigma}^2 - 20 \Sigma \ln \hat{\sigma}_i^2 = 123.96$$

در هر صورت هر دو مقدار از $\chi_4^2 \ {}^{\alpha=0.99}$ بزرگتر هستند.

چنانچه با فرض وجود واریانس ناهمسانی تخمین بزنیم داریم:

FGLS

$$I_t = -36.25 + 0.094 F_t + 0.33 C_t$$
$$(6.12) \quad (0.0007) \quad (0.030)$$

ML

$$I_t = -23.25 + 0.0943 F_t + 0.333 C_t$$
$$(4.815) \quad (0.0006) \quad (0.022)$$

فصل یازدهم

۱۱-۱ خود همبستگی[۴۵]

در داده های سری زمانی، مشکل شایع خود همبستگی یا همبستگی پیاپی است که بر روی جملات اخلال در دوره های مختلف وجود دارد.

برای مثال به ترسیم، پسماندهای حداقل مربعات معمولی در مثال زیر که از رگرسیون سرمایه گذاری حقیقی بر روی GNP حقیقی و نرخ بهره می باشد، توجه شود. یکی از دلایل وجود همبستگی پیاپی یا خود همبستگی این است که عوامل موثر از سری زمانی حذف شده است، عواملی که مانند متغیرهای موجود در مدل مهم و با یکدیگر ارتباط دارند.

دلیل دیگر برای منبع خود همبستگی، نحوه تولید داده های آماری است. برای مثال، در تعدیلات فصلی که روی متغیرها صورت می‌گیرد مانند CPI و GNP، مراکز تولید کننده‌های آماری در داخل داده هایشان خود همبستگی وارد می‌کنند.

مثال ۱

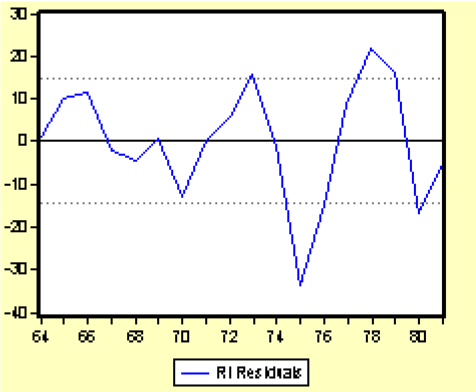
جدول (۱)

متغیر وابسته: سرمایه گذاری حقیقی		
متغیر های توضیحی	**ضرایب**	**انحراف معیار**
مقدارثابت	-۱۲/۵۳۳۶	۲۴/۹۲۰
نرخ بهره حقیقی	-۱/۰۰۱۴۴	۲/۳۶۹
GNPحقیقی	۰/۱۶۹۱۴	۰/۰۲۰۵
R^2	۰/۸۱۴	T=15

45. Autocottelation

منحنی پسماند رگرسیون تابع

تخمین سرمایه گذاری واقعی



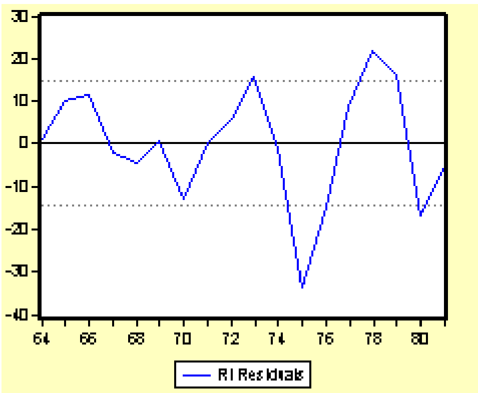
شکل (۱)

ترسیم جملات پسماند فوق نشان می‌دهد، اطلاع از علامت پسماند در یک دوره، نشان مناسبی از علامت پسماند در دوره بعد است. این بدین معنی است که، حداقل بخشی از اثر یک جمله اخلال، بر روی دوره‌های بعدی منتقل می‌شود. چنانچه جمله اخلال مثبتی، جمله اخلال مثبتی را بدنبال داشته باشد و یا جمله اخلال منفی را بدنبال داشته باشد، همبستگی مثبت داریم و چنانچه جمله اخلال مثبت جمله اخلال منفی و جمله اخلال منفی جمله اخلال مثبتی را به دنبال داشته باشد، همبستگی از نوع منفی است.

مثال ۲

نمودار پسماندهای معادله سرمایه گذاری به ما میگوید که علامت پسماند در یک دوره علامت خوبی از حرکت پسماند دوره بعد است. این اطلاعات به ما می‌گوید که اثر یک اختلال داده شده حداقل بطور جزیی در طول دوره ها منتقل می‌شود.

تخمین سرمایه گذاری واقعی



رگرسیون بر روی GNP حقیقی و نرخ بهره واقعی

مثال ۳

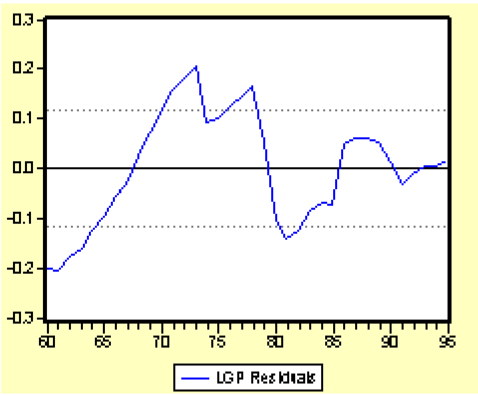
شواهد نشان میدهد که یک مدل رگرسیونی برای بازار بنزین امریکا (log(g/pop)حداقل باید شامل عرض از مبدا و (log (y/pop , log pg باشد. سایر متغیرها مانند قیمت و روند زمانی هم دارای حداقل اثری میباشند.

بعلاوه Chow test نشان دهنده روی دادن تغییرات ساختاری در ۱۹۷۴می‌باشد. در زیر نمودار پسماندها در حالتهای مختلف دیده میشود:

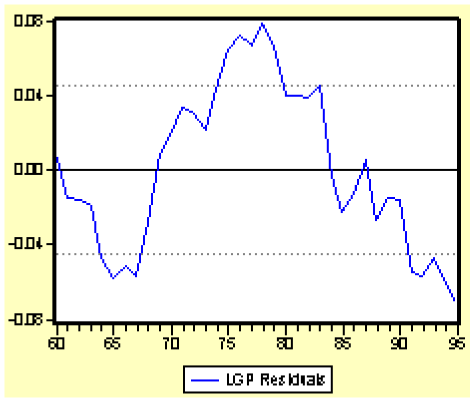
قسمت اول تا سوم می‌گویند زمانی که اجزای رگرسیون گسترش می‌یابند خود همبستگی پسماندها کاهش می‌یابد.

قسمت سوم مدلی را نشان می‌دهد که در آن همه ضرایب در معادله قبل و بعد از تغییرات ساختاری یکسان هستند.

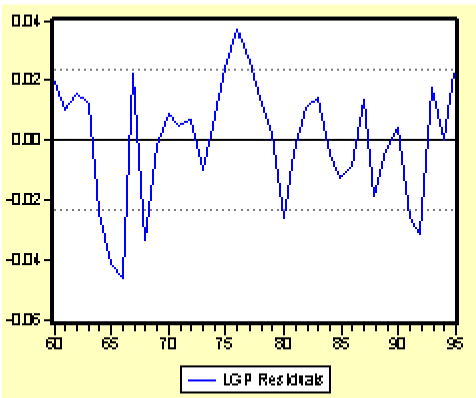
در قسمت آخر پسماندها با توجه به مدل غیر مقید با ضرایب متفاوت در دو دوره ۱۹۶۰- ۱۹۷۴ و۱۹۷۴- ۱۹۹۵ آورده شده از این دسته از پسماندها غیر همبسته اند. یعنی دامنه تغییرات کاهش می یابد.



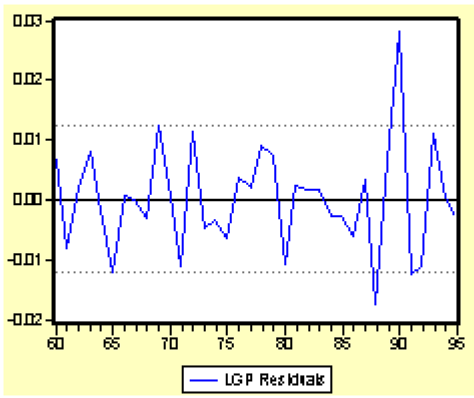
1.Regression on log(pg)



2.Regression on log(pg),log(y/pop)



3.Full Regression



4.Full regression, separate coefficients

۱۱-۲ تخمین تحت شرایط همبستگی پیاپی

مدل $Y = X\beta + U$ را در نظر بگیریم و فرض کنیم u از نوع AR (1) می‌باشد، بطوریکه:

$$u_t = \rho u_{t-1} + \varepsilon_t$$

بطوریکه:

$$E(\varepsilon_t) = 0 \quad , \quad E(\varepsilon_t^2) = \sigma^2 \; , \quad E(\varepsilon_t \, \varepsilon_s) = 0 \quad t \neq s$$

در اینجا علاوه بر White Noise[۴۶] بودن جمله ε_t در فراگرد AR (1)، شرط $|\rho| < 1$ را نیز لحاظ می کنیم. بدین ترتیب فراگرد تصادفی از نوع مانا می باشد. برای محاسبه ماتریس، واریانس – کوواریانس uu' می‌توان به ترتیب زیر عمل نمود:

$$Var\ (u_t) = \rho^2 Var(u_{t-1}) + Var\,(\varepsilon_t)$$

$$\sigma_u^2 = \frac{\sigma_\varepsilon^2}{1-\rho^2}$$

$$E\,(u_t u_{t-1}) = \rho \frac{\sigma_\varepsilon^2}{1-\rho^2}$$

$$E\,(u_t u_{t-h}) = \rho^h \frac{\sigma_\varepsilon^2}{1-\rho^2}$$

$$\therefore E\,(UU') = \frac{\sigma_\varepsilon^2}{1-\rho^2} \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^{T-1} \\ \rho & & \rho & \dots & \rho^{T-2} \\ & & & & \vdots \\ & & & & \rho \\ & & \dots & \rho & 1 \end{bmatrix} = \sigma_\varepsilon^2 W$$

چنانچه مقدار ρ معلوم باشد، TGLS بکار می بریم یا در مدل تبدیل یافته، OLS بکار می‌بریم.

$$W^{-1} = \begin{bmatrix} 1 & -\rho & 0 & 0 & \dots & 0 \\ -\rho & 1+\rho^2 & -\rho & 0 & \dots & 0 \\ 0 & -\rho & 1+\rho^2 & -\rho & \dots & 0 \\ 0 & 0 & -\rho & 1+\rho^2 & \dots & 0 \\ \vdots & & & \ddots & 1+\rho^2 & -\rho \\ 0 & & & & -\rho & 1 \end{bmatrix} = P'P$$

۴۶. اگر فراگر تصادفی دارای میانگین صفر و واریانس σ^2 باشد و با یکدیگر وابستگی نداشته باشد آنرا white Noise گویند.

$$P=\begin{bmatrix} \sqrt{1-\rho^2} & 0 & 0 & \dots & 0 \\ -\rho & 1 & 0 & \dots & 0 \\ 0 & -\rho & 1 & \dots & \vdots \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & -\rho & 1 \end{bmatrix}$$

مدل تبدیل یافته شده عبارتست از:

$$PY = PX\beta + PU$$
$$Y_* = X_* \beta + U_*$$
$$t=1 \quad \sqrt{1-\rho^2}\ y_1 = \sqrt{1-\rho^2}\ X_{11}\beta_1 + \sqrt{1-\rho^2}\ X_{12}\beta_2 + \dots + \sqrt{1-\rho^2}\ X_{1k}\beta_k$$
$$t=2 \quad (y_2-\rho y_1) = (X_{21}-\rho X11)\beta_1 + (X_{22}-\rho X12)\beta_2 + \dots + (X_{2k}-\rho X_{1k})\beta_k$$
$$t\ge 2 \ y_t - \rho y_{t-1} = (X_{t1}-\rho X_{t-1})\beta_1 + \dots + (X_{tk}-\rho X_{t-1,k})\beta_k + u_t - \rho u_{t-1}$$

و چنانچه بر روی ستونهای X مدل را بنویسیم داریم:

$$t\ge 2 \qquad y_t = X_{t0}\beta + u_t$$
$$\rho y_{t-1} = \rho X_{t-1,0}\beta + \rho u_{t-1}$$
$$\overline{}$$
$$y_t - \rho_{t-1} = (X_{t0} - \rho X_{t-1})\beta + u_t - \rho u_{t-1}$$

می توانیم بدون آنکه از TGLS اطلاعی داشته باشیم، با از دست دادن مشاهده اول، چنانچه ρ معلوم باشد، به روش OLS، β را تخمین بزنیم. اما ρ معمولا معلوم نیست. لذا می توانیم ρ را از مدل AR(1) تخمین بزنیم و در مدل قرار دهیم و سپس β را برآورد نمائیم.

۱۱-۳ روش تخمین به روش حداکثر درست نمایی

تابع حداکثر درست نمایی برای جمله u درشرایطی که $E(uu')=\sigma^2 W$ می باشد، بصورت زیر است:

$$L(Y;\rho,\beta,\sigma_\varepsilon^2) = (2_\pi)^{-\frac{T}{2}} 0| \sigma_\varepsilon^2 W |^{-\frac{T}{2}} \exp\left[-\frac{1}{2}(Y-X\beta)'(\sigma_\varepsilon{}^2 W)^{-1}(Y-X\beta)\right]$$

با گرفتن لگاریتم از طرفین رابطه فوق داریم:

$$Log\,L(Y;\rho,\beta,\sigma_\varepsilon^2)=Log\left[2_\pi^{-\frac{T}{2}}\,|\,\sigma_\varepsilon^2 W\,|^{-\frac{1}{2}}\,\exp\left[-\frac{1}{2}(Y-X\beta)'(\sigma_\varepsilon{}^2 W)^{-1}(Y-X\beta)\right]\right]$$

$$=-\frac{T}{2}Log\,2\pi-\frac{T}{2}Log\,\sigma_\varepsilon^2-\frac{1}{2}Log\,|\,W\,|-\frac{1}{2\sigma_\varepsilon^2}(y-X\beta)'\,w^{-1}(y-X\beta)$$

چنانچه مقدار ثابت را حذف کنیم و در عدد ۲- ضرب کنیم داریم:

$$Log\,L(Y;\rho,\beta,\sigma_\varepsilon^2)=TLog\sigma_\varepsilon^4\,|\,W\,|+\frac{1}{\sigma_\varepsilon^2}(Y-X\beta)'\,(W)^{-1}(Y-X_\beta)\qquad(۱)$$

حال تابع فوق را ابتدا نسبت به σ_ε^2 ، مینیمم می کنیم و $\hat{\sigma}_\varepsilon^2$ را بصورت زیر بدست می آوریم:

$$\hat{\sigma}_\varepsilon^2=\frac{1}{T}(Y-X\beta)'W^{-1}(Y-X\beta)$$

چنانچه $\hat{\sigma}_\varepsilon^2$ را در رابطه (۱) قرار دهیم داریم:

$$Log\,L(Y;_{\rho,\beta})=T\,Log\quad\hat{\sigma}_\varepsilon^2+Log\,|\,W\,|+T$$

$$=TLog\frac{1}{T}+T\,Log\,(Y-X_\beta)+Log\,|\,W\,|+T$$

T و جمله اول را حذف و طرفین را به T تقسیم می کنیم، خواهیم داشت:

(۲)

$$Log\;\,L^*(Y;\rho,\beta)=Log\,(Y-X\beta)'W^{-1}(Y-X\beta)|\,W\,|^{\frac{1}{T}}$$

تابع فوق را نسبت به ρ و β مینیمم می کنیم.

از طرفی داریم:

$$W^{-1}=P'P$$

$$|\,W^{-1}|=|P'P\,|=|P'\,|\;|\,P\,|=1-\rho^2$$

چون: $\qquad |\,P\,|=\sqrt{1-\rho^2}$

$$\Rightarrow|\,W\,|=\frac{1}{1-\rho^2}$$

لذا رابطه (۲) را می توانیم بصورت زیر بنویسیم:

$$L^{*}(Y;\rho,\beta)=(1-\rho^2)^{-\frac{1}{T}}(Y-X\beta)'P'P(Y-X\beta)$$
$$=(1-\rho^2)^{-\frac{1}{T}}(Y_*-X_*\beta)'(Y_*-X_*\beta)$$
$$=(1-\rho^2)^{-\frac{1}{T}}\left[\sum_{t=1}^{T}(Y_{*.t}-X_{*.t}.\beta)^2\right]$$
$$=(1-\rho^2)^{-\frac{1}{T}}\left\{(1-\rho^2)(Y_1-X_{10}\beta)^2+\sum_{t=2}^{T}[(Y_t-\rho y_{t-1})-(X_{t0}-\rho X_{t-1,0})\beta]^2\right\}$$

اگرتعداد مشاهدات باندازه کافی بزرگ شود داریم:

$$Lim_{T\to\infty}(1-\rho^2)^{-\frac{1}{T}}=1$$

و بالاخره می توانیم از مشاهده اول و در نتیجه $(1-\rho^2)(Y_1-X_{10}\beta)^2$ صرفنظر نمائیم و تابع زیر را مینیمم کنیم:

(۳)

$$L^{***}(\rho,\beta)=\sum_{t=2}^{T}[(Y_t-\rho Y_{t-1})-(X_{t0}-\rho X_{t-1,0})\beta]^2$$

اگر اندازه نمونه باندازه کافی بزرگ است، تخمین ρ و β در فرم درجه دوم فوق تخمین زننده هایی نزدیک به MLE می باشند.

یـــــــــادآوری:

۱- اگر ρ نزدیک به یک باشد، تقریب نمودن رابطه (۲) و (۳) شاید صحیح نباشد.

۲- تخمین ρ و β که با مینیمم کردن L^{***} بدست می آیند بطور مجانبی با $\hat{\rho}_{ML},\hat{\beta}_{ML}$ یکسان می باشند.

۳- در نمونه های کوچک اگر چنانچه مدل (۳) را ابتدا نسبت به β ثابت نگه داریم و ρ را بدست بیاوریم و یا برعکس ابتدا نسبت به β تخمین بزنیم و سپس نسبت به ρ، شاید متفاوت باشد. رابطه (۳) را می توان بصورت زیر نیز نوشت:

(۴)

$$LogL^{**}(\rho,\beta)=\sum_{t=2}^{T}[(Y_t-X_{t0}\beta)-\rho(Y_{t-1}-X_{t-1,0}\beta)]^2$$

۱۱-۴ مراحل روش تخمین کوکران اورکات[۴۷]:

۱- Yرا بر روی X رگرس می کنیم، یعنی بجای $\rho=0$ قرار می‌دهیم و رابطه (۴) را نسبت به β مینیمم می‌کنیم. پس از بدست آوردن $\hat{u}_{ols}, \hat{\beta}_{ols}$ استخراج و سپس $\hat{u}_t$ را بر روی $\hat{u}_{t-1}$ برای بدست آوردن $\hat{\rho}$ بکار می بریم.

$$\hat{\rho}=\frac{\sum_{t=2}\hat{u}_t\hat{u}_{t-1}}{\sum_{t=2}\hat{u}_t^2} \qquad ; \qquad Min LogL^{**}(\hat{\rho}_{ols,}\,\beta)$$

۲- $\hat{\rho}_{ols}$ را داخل رابطه (۴) قرار می دهیم و آنرا نسبت به βمینیمم می‌کنیم و $\hat{\beta}_{co}$ را بدست می‌آوریم.

۳- $LogL^{**}(\rho, \hat{\beta}_{co})$ را مجددا نسبت به ρ مینیمم می‌کنیم تا $\hat{\rho}_{co}$بدست آید و می‌توانیم این روش را تکرار کنیم.

یـــــادآوری:

۱- اگر مرحله یک را با یک تخمین زننده سازگار برای βآغاز کنیم و آنرا برای چند بار تکرار کنیم، آنگاه نتایج بدست آمده برای β و ρ سازگار است و بطور مجانبی با MlE یکسان خواهد بود.

۲- برای T ثابت، اگر روش کوکران – اورکات را تکرار کنیم، همگرا خواهد شد ولی ممکن است بیش ازیک نقطه مینیمم برای رابطه (۴) پیدا کنیم، از این رو روش تکراری ممکن است بستگی به مقادیر اولیه داشته باشد.

۱۱-۵ روش و مراحل کار Hilderth – LU

۱- مقداری بین ۱ و ۱- برای ρدر نظر می‌گیریم.

۲- برای هر مقدار از ρ تابع (۴) را نسبت به βمینیمم می‌کنیم.

$$\Rightarrow \hat{\beta}(\rho) \Rightarrow RSS(\rho) = LogL^{**}(\rho, \hat{\beta}(\rho))$$

۳- مقداری برای ρ انتخاب می کنیم که $RSS_{(\rho)}$ را مینیمم کند.

$$\Rightarrow \hat{\rho}, \hat{\beta}(\hat{\rho})$$

۴- این روش را با تکرار و انتخاب مقادیر محدود در حوالی $\hat{\rho}$ادامه می‌دهیم. هر زمان که بین $\hat{\beta}$ تغییرات تثبیت شد، متوقف می‌نمائیم.

47. Cochran - Orcutt

۱۱-۶ روش و مراحل کار دوربین[48]

رابطه (۳) را می توانیم بصورت زیر بنویسیم:

$$\sum_{t=2}^{T}[Y_t-(\rho Y_{t-1}+X_{t0}\beta-\rho X_{t-1,0}\beta)]^2 \qquad (۵)$$

چنانچه رابطه بین ρ و β را نادیده بگیریم، رابطه (۵) تابع هدف، OLS خواهد بود. یعنی Y_t را بر روی $-X_{t-1,0},X_{t0},Y_{t-1}$، رگرس می کنیم، ρ ضریب متغیر Y_{t-1} است. اما اگر به رابطه بین ρ و β توجه کنیم و $\gamma=\rho\beta$ را به روش NLS تخمین بزنیم. $\hat{\gamma}$ بطور سازگار برآورد می شود. آنگاه از $\hat{\rho}$ برای FGLS استفاده می کنیم.

خلاصه

مدل $Y=X\beta+U$ را با توجه به $u_t=\rho u_{t-1}+\varepsilon_t$ در نظر بگیریم:

$|\rho|<1 \quad , \varepsilon_t \quad i.i.d$

۱۱-۷ روش FGLS

۱- مقدار اولیه ای برای $\hat{\rho}$ از روش (Cochran – Orcutt , Hilderth – LU, Durbin) بدست می آوریم.

۲- $\hat{\beta}_{FGLS}=(X'(W(\hat{\rho}))^{-1}X)^{-1}(X'(W(\hat{\rho}))^{-1}y$

یـــــادآوری:

اگر داده های آماری فصلی باشد، و رابطه بصورت زیر:

$u_t=\rho u_{t-4}+\varepsilon_t$

سپس، GLS را محاسبه می کنیم، مدل تبدیل یافته بصورت زیر است:

$$\sqrt{1\rho^2}Y_t=\sqrt{1-\rho^2}\;X_{t0}\beta+\sqrt{1-\rho^2}u_t \qquad t=1,2,3,4$$

$$y_t-\rho y_{t-4}=(X_{t0}-\rho X_{t-4})\;\beta+u_t-\rho u_{t-4} \qquad t=5,6,\dots,T$$

نحوه ادامه کار برای رسیدن به MLE, FGLS همانند مدل AR(1) می باشد.

۱۱-۸ آزمون برای تشخیص همبستگی پیاپی (خود همبستگی)

48. Durbin.

در اینجا می خواهیم فرضیه H_0 را بصورت: $H_0: W=I$ بررسی کنیم:

۱۱-۸-۱ آزمون دوربین واتسون[۴۹]:

آماره d را بصورت زیر تعریف می‌کنیم:

$$d=\frac{\sum_{t=2}^{T}[\hat{u}_t-\hat{u}_{t-1}]^2}{\sum_{T=1}^{T}\hat{u}_1^2}$$

$$\hat{u}_t=y_t-X_{t0}\,\hat{\beta}_{ols}$$

$$\hat{U}'\hat{U}=U'MU \quad \Rightarrow d=\frac{U'AU}{U'MU}$$

بطوریکه:

$$M=I-X\,(X'X)\,X'$$

$$A=\begin{bmatrix} 1 & -1 & 0 & \dots \\ -1 & 2 & -1 & \dots \\ 0 & -1 & 2 & \dots \\ \vdots & \vdots & \ddots 2 & -1 \\ 0 & 0 & -1 & 1 \end{bmatrix}$$

u_t یک فراگرد مانا است و $\rho_1=E(u_t u_{t-1})$ است، در صورتیکه ارتباط بین u_t و u_{t-1} ، کامل باشد $\rho_1=1$ است و لذا $\hat{u}_t=\hat{u}_{t-1}$ خواهد شد. سپس d = 0 خواهد شد.

چنانچه $\rho_1=0$ باشد در آنصورت

$$d\approx\frac{\frac{1}{T}\sum(\hat{u}_t^2-\frac{1}{T}2\sum\hat{u}_t\,\hat{u}_{t-1}+\frac{1}{T}\sum\hat{u}_{t-1}^2)}{\frac{1}{T}\sum\hat{u}_t^2}$$

چنانچه T باندازه کافی بزرگ باشد:

$$d=\frac{\sigma_u^2+0+\sigma_u^2}{\sigma_u^2}=2$$

چنانچه $\rho_1=-1$ باشد، در آنصورت:

49. Durbin - Watson

$$u_t = -u_{t-1}$$
$$\Rightarrow \hat{u}_t = -\hat{u}_{t-1}$$
$$\Rightarrow d = \frac{4\sum \hat{u}_t^2}{\sum \hat{u}_t^2} = 4$$

روشن است که $0 \le d \le 4$ است.

دلیل این دامنه به لحاظ رابطه زیر است:

$$d = \frac{\sum\limits_{t=2} \hat{u}_t^2 - 2\sum \hat{u}_t \hat{u}_{t-1} + \sum\limits_{t=2} \hat{u}_{t-1}^2}{\sum \hat{u}_t^2}$$

جمله اول بزرگتر از یک و جمله سوم کوچکتر از یک است

پس:

$$d \le 2 - 2\frac{\sum\limits_{t=2} \hat{u}_t \hat{u}_{t-1}}{\sum\limits_{t=1} \hat{u}_t^2} = 2 - 2\hat{\rho}$$

برای آنکه نشان دهیم $d \le 4$ است باید رابطه زیر را ثابت کنیم:

$$\left| \sum_{t=2}^{T} \hat{u}_t \hat{u}_{t-1} \right| \le \sqrt{\sum_{t=2} \hat{u}_t^2} \sqrt{\sum_{t=2}^{T} \hat{u}_{t-1}^2}$$
$$d \le 2 + 2\frac{\left| \sum\limits_{t=2} \hat{u}_t \, \hat{u}_{t-1} \right|}{\sum\limits_{t=1} \hat{u}_t^2} \le 2 + 2\frac{\sqrt{\sum\limits_{t=2} \hat{u}_t^2} \quad \sqrt{\sum\limits_{t=2} \hat{u}_{t-2}^2}}{\sum\limits_{t=1} \hat{u}_t^2}$$
$$d \le 2 + 2 = 4$$

لذا نشان داده شده، d همواره بین صفر و ۴ قرار دارد. چنانچه پسماندها همبستگی پیاپی مثبت داشته باشند، مقدار محاسبه شده d کوچک خواهد بود و اگر پسماندها همبستگی پیاپی منفی داشته باشند مقدار محاسبه شده d بزرگ خواهد بود. البته می توان با استفاده از مقادیر مربوط به جدول برای d_l و d_u آزمون فرضیه را دقیق تر انجام داد.

فصل دوازدهم

۱-۱۲ هم خطی[۵۰]

در تحلیل‌های مدل رگرسیون فرض می‌کردیم که ما تریس مشاهدات برروی متغیرهای مستقل یعنی ماتریس X دارای مرتبه کامل باشد. یعنی رابطه خطی کامل بین متغیرهای مستقل وجود نداشته باشد. می‌دانیم برای آنکه معکوس $(X'X)$ وجود داشته باشد، X باید دارای مرتبه ستونی کامل باشد. موردی را در نظر بگیریم که فقط دو متغیر توضیحی و عرض از مبدأ داریم. برای هر کدام از شیب‌های مسئله داریم:

$$\operatorname{Var}(\beta_k)=\frac{\sigma^2}{(1-r_{12}^2)\sum_{i=1}^{T}(x_{ik}-\bar{x}_k)^2}=\frac{\sigma^2}{(1-r_{12}^2)\,S_{kk}} \qquad k=1,2 \qquad (۱)$$

اگر دو متغیر کاملاً به یکدیگر وابسته هستند، واریانس β_k به سمت بینهایت میل می‌کند. وجود ارتباط خطی کامل در بین رگرسورها نقض جدی فروض مدل است. مورد شایعی که وجود دارد این است که ارتباط بین دو متغیر خیلی بالاست اما کامل نیست. در این حالت مدل رگرسیون دارای کلیه ویژگی‌های خود می‌باشد هر چند بالقوه دارای مسایل آماری حادی می‌باشد.

۲-۱۲ هم خطی کامل[51]

فرض می‌کنیم تابع مصرف را بصورت زیر تصریح نموده باشیم:

---

50 - Multicollinearity.
51 -Perfect Collinearity.

$C= \beta_1 + \beta_2$ (درآمد از طریق غیرکار) + β_3 (حقوق) + β_4 (کل درآمد) + ε

در تابع فوق امکان تفکیک اثرات هر کدام از اجزاء درآمد برروی مصرف وجود ندارد. مثلاً اگر تابع فوق را بصورت زیر بنویسیم:

$$E(C)=\beta_1+\beta_2 N+\beta_3 S+\beta_4 T \quad ; \quad T=N+S$$

گیریم q مقدار غیرصفری باشد که پارامترهای مدل را بصورت زیر بازنویسی کنیم:

$$\beta'_2=\beta_2+q \quad , \quad \beta'_3=\beta_3+q \quad , \quad \beta'_4=\beta_4-q$$

لذا داریم:

$$E(C)=\beta_1+\beta'_2 N+\beta'_3 S+\beta'_4 T$$

و این مدل برای پارامترهای مختلف قابل تصریح است.

برخلاف مرتبه کامل این مدل اجازه می‌دهد که مقدار E(C) برای مقادیر مختلف پارامترها یکسان باشد.

البته در اینجا باید توجه کنیم که مشکل داده آماری نداریم بلکه مشکل در تصریح مدل است. اصطلاحاً می‌گویند پارامترهای این مدل قابل شناسایی نیستند و این مشکل را تحت عنوان مسئله تشخیص یا شناسایی بررسی خواهیم کرد. در این مدل ترکیبات بی‌شماری از پارامترها می‌توانیم داشته باشیم که میانگین انتظاری C یکسان است. لذا مهم نیست که چه مقدار داده آماری توانسته‌ایم گرد آوریم، پارامترها قابل تخمین نخواهند بود. در نتیجه فرض مرتبه کامل بعنوان فرض تشخیص می‌تواند در نظر گرفته شود، باین معنی که برای بردار معین x، فقط یک دسته پارامتر می‌باید سازگار با مقدار معین E(y) باشد.

مثال ۱:

$$Y=\beta_1 x_1+\beta_2 x_2+u$$
$$x_2=a x_1$$
$$X'X=\sum x_2^2\begin{bmatrix}1 & \alpha\\ \alpha & \alpha^2\end{bmatrix}$$
$$|X'X|=0$$
$$R(X'X)=1$$

لذا پارامترهای β_1و $\beta 2$ را نمی‌توانیم بصورت منحصر بفرد باروش OLSتخمین زده وبدست آوریم.

معادلات نرمال
$$\begin{cases}\sum x_2^2(\hat\beta_1+\alpha\hat\beta_2)=\sum x_2 y\\ \alpha\sum x_2^2(\hat\beta_1+\alpha\hat\beta_2)=\alpha\sum x_2 y\end{cases}$$

دو معادله تبدیل به یک معادله می‌شود بصورت زیر:

$$\hat\beta_1+\alpha\hat\beta_2=\frac{\sum x_2 y}{\sum x_2^2}$$
$$\Rightarrow E(y)=(\hat\beta_1+\alpha\hat\beta_2)x_2=yx_2$$

لذا E(y) می‌تواند بطور منصر بفرد بدست آید چون γ را می‌توانیم بصورت منحصر بفرد بدست آوریم.

۳-۱۲ هم خطی ناقص

مثال ۲:

وقتی مرتبه ماتریس کوچکتر از K باشد مثلاً $L<K$ ،درآنصورت L پارامتر قابل تخمین خواهند بود. اهمیت این درآنست که در K ستون فقط L منبع تغییر وجود دارد. لذا داده های آماری فقط اجازه تخمین L پارامتر را بما می‌دهد. بیشتر شایع است که متغیرها با یکدیگر رابطه داشته باشند ولی نه رابطه‌ای کامل، مورد هم خطی ناقص یا همبستگی بالا بین متغییرها، برخلاف هم خطی کامل یک مشکل در داده‌های آماری است.

مشکل تخمین ازنوع عدم تشخیص نخواهد بود اما به دقت تخمین زننده ها برمی‌گردد. هر چقدر رابطه بین رگرسورها بیشتر باشد، دقت تخمین زننده‌ها نیز کاهش می‌یابد. این مسئله رادر مدل زیر بررسی می‌کنیم. مدل $Y=X\beta+U$ را در نظر بگیریم. بطوریکه $X=[W \quad X_k]$ ، درنتیجه داریم:

$$X=[W \quad x_k]\beta+U$$

اثرات هم خطی ناقص را بر روی دقت تخمین زننده ها بررسی می‌کنیم. رابطه (۱) را می‌توانیم بصورت زیر بنویسیم:

$$(۲) \quad Var(\hat{\beta}_k)=\frac{\sigma^2}{\hat{V}_k'\hat{V}_k}$$

بطوریکه:

$$\hat{V}_k=(I-W(W'W)^{-1}W')X_k$$

V_k عبارتست از جمله پسماند ناشی از رگرسیون X_k بر روی سایر رگرسور ها اگر x_k در حد بالایی با سایر رگرسورها همبسته باشد، آنگاه وقتی $\hat{V}_k'\hat{V}_k$ کوچک می‌شود، چون مجموع مجذورات خطا همراه باافزایش هم خطی بین متغییر K ام وسایر متغییرها کاهش پیدا می‌کند. درنتیجه واریانس $\hat{\beta}_k$ افزایش پیدا می‌کند. ولی توجه داشته باشیم، همه پارامترها شبیه بهم وبطور یکسان ازاین قاعده متاثر نمی‌شوند.

به مثال عددی زیرتوجه کنیم:

$$\begin{array}{ccccc} X'X & (X'X)^{-1} & R^2_{1,23} & R^2_{2,13} & R^2_{3,12} \\ \begin{bmatrix} 10 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix} & \begin{bmatrix} 0.1 & 0 & 0 \\ 0 & 0.2 & 0 \\ 0 & 0 & 1 \end{bmatrix} & 0 & 0 & 0 \end{array}$$

در اینجا متغییرهای مستقل کاملاً بر هم عمودند . واریانس $\hat{\beta}_3$ بزرگتر از $\hat{\beta}_2$ و بزرگتر از $\hat{\beta}_1$ است.

واریانس $\hat{\beta}_3$ ده برابر واریانس $\hat{\beta}_1$ است ودترمینان $X'X$ برابر ۱۰ می‌باشد.

$$\begin{bmatrix} 10 & 5 & 2 \\ 5 & 5 & 2 \\ 2 & 2 & 1 \end{bmatrix} \begin{bmatrix} 0.2 & -0.2 & 0 \\ -0.2 & 1.2 & -2 \\ 0 & -2 & 2 \end{bmatrix} \quad 0.50 \quad 0.83 \quad 0.80$$

ملاحظه میشود هرچند واریانس $\hat{\beta}_3$، ۲۵برابر $\hat{\beta}_1$است. دترمینان $X'X$ برابر ۵ است.

$$\begin{bmatrix} 10 & 6 & 3 \\ 6 & 5 & 2 \\ 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & -2 \\ -3 & -2 & 14 \end{bmatrix} \quad 0.90 \quad 0.80 \quad 0.92$$

در این حالت واریانس $\hat{\beta}_1$ بالا رفته و برابر واریانس $\hat{\beta}_2$ شده است. دترمینال $X'X$ برابر یک می‌باشد.

$$\begin{bmatrix} 10 & 7 & 1.5 \\ 7 & 5 & 1 \\ 1.5 & 1 & 1 \end{bmatrix} \begin{bmatrix} 5.3 & -7.3 & -9.7 \\ -7.3 & 10.3 & 0.7 \\ -0.7 & 0.7 & 1.3 \end{bmatrix} \qquad 0.981 \quad 0.980 \quad 0.25$$

در این وضعیت واریانس $\hat{\beta}_3$ کوچکترین واریانس شده است و خیلی بزرگتر از واریانس $\hat{\beta}_3$ در حالت اول نیست. درحالیکه واریانس $\hat{\beta}_3$ و $\hat{\beta}_2$ حدوداً ۵۰ برابر از حالت اول بزرگتر شده‌اند. دترمینال $\hat{\beta}_2$ برابر ۰/۷۵ است. پس نتیجه این است که اثرات همخطی ناقص بر روی ضرایب مختلف متفاوت است.

۴-۱۲ تشخیص هم خطی

در مثال عددی فوق اگر به R^2 ها توجه شود، بین آنها و افزایش واریانس‌ها نسبت به حالت اول رابطه ای وجود دارد. چنانچه فرض کنیم TSS_k ، کل مجذورات خطا برای انحراف از میانگین X_k باشد.

$RSS_k = \hat{V}'_k \hat{V}_k$ ، مجموع مجذورات خطا وقتی X_k روی W رگرس می‌شود محسوب می‌گردد، در آنصورت: مجذور ضریب همبستگی

$$R_k^2 = 1 - \frac{RSS_k}{TSS_k}$$

یا :

$$RSS_k = TSS_k\,(1 - R_k^2)$$

از رابطه (۲) داریم:

$$Var(\hat{\beta}_k) = \frac{\sigma^2}{RSS_k} = \frac{\sigma^2}{TSS_k\,(1-R_k^2)}$$

اگر فرض کنیم $\hat{\beta}_0$ عبارتست از تخمین پارامترها در حالتی که $X'X$ ماتریس قطری باشد. آنگاه $R_k^2 = 0$ است و واریانس $\hat{\beta}_0$ برابر است با $\frac{\sigma^2}{TSS_k}$.

لذا می‌توانیم از نسبت این دو واریانس بعنوان یک شاخص برای اندازه‌گیری بزرگی واریانس‌های ضرایب مختلف استفاده کنیم.

$$\frac{Var(\hat{\beta}_k)}{Var(\hat{\beta}_0)} = \frac{1}{1-R_k^2}$$

مـثـال عـددي :

R_k^2 :	0.5	0.8	0.9	0.95	0.96	0.97	0.98	0.99	0.999
$\frac{Var(\hat{\beta}_k)}{Var(\hat{\beta}_0)} =$	2	5	10	20	25	33	50	100	1000

روش جایگزین برای تشخیص هم خطی استفاده از عدد وضعیت[۵۲] ماتریس $X'X$ است. این روش توسط Belsley et al (1980) با استفاده از مقادیر ویژه ماتریس بکار برده شده است.

عدد وضعیت یک ماتریس جذر نسبت بزرگترین مقدار ویژه یک ماتریس به کوچکترین مقدار ویژه[۵۳] آن ماتریس بدست می آید.

52 -Condition Number.

۵۳- مقادیر ویژه ماتریس در بخش ماتریسها آمده است

دترمینال یک ماتریس را می‌توان برحسب حاصلضرب مقادیر ویژه[۵۴] یک ماتریس (که معمولاً با نماد λ نشان می‌دهند) نشان داد یعنی:

$$|X'X| = \lambda_1 \lambda_2 \ldots \lambda_k$$

دترمینال خیلی کوچک باین معنی است که بعضی یا خیلی از مقادیر ویژه کوچک هستند.

$$\gamma = \text{ConditionNumber} = \left(\frac{\lambda \max}{\lambda \min}\right)^{1/2}$$

اگر X_k ها همه بر هم عمود باشند و اصلاً رابطه‌ای بین آنها وجود نداشته باشد γ=۱ است[۵۵] و این در حالی است که R_k^2 ها برابر صفر است.

هر چقدر همبستگی متغیرها با یکدیگر بیشتر باشد، γ نیز افزایش می‌یابد. $\gamma > ۲۰$، علامت خطر برای وجود هم خطی است.

مثال ۲:

فرض کنیم در مثال فوق α=۲ و $X'Y = \begin{bmatrix} 5 \\ 10 \end{bmatrix}$ و $X'X = \begin{bmatrix} 10 & 20 \\ 20 & 40 \end{bmatrix}$

حل:

$$|X'X| = 0$$

$$10\,\hat{\beta}_1 + 20\hat{\beta}_2 = 5$$

$$20\hat{\beta}_1 + 40\hat{\beta}_2 = 10$$

نشان داده می‌شود که برای مقادیر مختلف $\hat{\beta}_1$ و $\hat{\beta}_2$ می‌توانیم مقدار منحصر بفردی برای میانگین y بدست آوریم.

$\hat{\beta}_2$	$\hat{\beta}_1$	$\hat{\beta}_1 + \hat{\beta}_2$	$E(Y \mid X_1 = 20)$
0	0.5	0.5	10
1	−1.5	0.5	10
−1	2.5	0.5	10
−10	20.5	0.5	10

$$\hat{\beta}_1 = \frac{5 - 20\hat{\beta}_1}{10}$$

گرچه $\hat{\beta}_1$ و $\hat{\beta}_2$ را نمی‌توانیم بطور جداگانه تخمین بزنیم، ولی ترکیب خطی آنها قابل تخمین است و $E(y)$ برای هر مقدار را می‌توان تخمین زد.

۱۲-۴-۱ تجزیه واریانس ضرایب مدل رگرسیون[۵۶]

چنانچه A ماتریس بردارهای ویژه ماتریس $X'X$ باشد، آنگاه داریم:

$$A'X'XA = \Lambda = \text{diag}(\lambda_i)$$

به طوریکه: $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_k$

54 - Eigen Values.

۵۵- برای از بین بردن مقیاس هر ستون را بر $X_k' X_k$ تقسیم می‌کنیم.

56 - Regression Coefficient Decomposition Variable.

و سپس داریم: $X'X = A\Lambda A'$

در نتیجه: $(X'X) = A\Lambda A'$

یا: $(X'X)^{-1} = A\Lambda^{-1}A' = \sum_{i=1}^{k} \lambda_i^{-1} a_i a_i'$

بطوریکه a_i بردارهای ویژه متناظر بامقادیر ویژه هستند.

ماتریس واریانس- کواریانس پارامتر $\hat{\beta}$ رامی‌توانیم بصورت زیر بنویسیم:

$$VC(\hat{\beta}) = \sigma^2 (X'X)^{-1} = \sigma^2 A\Lambda^{-1}A'$$
$$= \sigma^2 \sum_{i=1}^{k} \lambda_i^{-1} a_i a_i'$$
$$Var(\hat{\beta}_j) = \sigma^2 \left(\frac{a_{j1}^2}{\lambda_1} + \frac{a_{j2}^2}{\lambda_2} + \dots + \frac{a_{jk}^2}{\lambda_k}\right)$$

این معادله می‌تواند بیانگر دقت تخمین $\hat{\beta}_j$ بوسیله OLS باشد.

عواملی که بر دقت تخمین زننده ها می‌تواند موثر باشند را بشرح زیر می‌توان بیان کرد:

۱- دقت تخمین هر پارامتر بصورت معکوس با σ^2 رابطه دارد. یعنی هرچه اغتشاش زیاد داشته باشیم، نتایج تخمین از اطمینان کمی برخوردارند.

۲- تغییرات متغیرهای مستقل بیان کننده مقادیر ویژه $X'X$ است که به دو صورت زیر بر روی تخمین زننده ها اثرمی‌گذارد.

الف- اگر تغییرات کل متغیرهای مستقل را با اثر ماتریس $(X'X) = \sum \lambda_i$ [۵۷] اندازه گیری کنیم، آنگاه افزایش درکل تغییرات، بشرط آنکه سایر عوامل ثابت باشد، دقتOLS را بالا می‌برد.

ب- مقادیر ریشه های ویژه از طریق تساوی و یا عدم تساوی بین خودشان روی دقت تخمین زننده ها اثر می‌گذارند. یعنی اگر λ_k نسبت به سایر λ_i ها خیلی کوچک است، آنگاه $\sigma^2 = \frac{a_{jk}^2}{\lambda_k}$ خیلی بزرگ می‌شود و لذا $Var(\hat{\beta}_j)$ ممکن است بزرگ شود.

وقتی رگرسورها با یکدیگر رابطه بالایی داشته باشند معمولاً موارد زیر را مشاهده می‌کنیم:

الف) تغییرات کوچکی در نمونه داده‌های آماری تغییرات زیادی را در تخمین پارامترها موجب می‌گردد.

ب) علیرغم $\hat{\beta}_j$ بالا که نشان از ارتباط بین متغیرهای مستقل و وابسته است ولی به لحاظ داشتن انحراف معیارهای بزرگ، تخمین ضرایب از سطح معنی‌داری برخوردار نیستند.

ج) علامت ضرایب خلاف انتظارات تئوریک است.

دو راه حل مکانیکی برای برخورد با مهمترین مسئله‌ای که هم خطی ایجاد می‌کند (یعنی واریانس‌های بزرگ) پیشنهاد شده است. راه حل اول استفاده از روش تخمین ریج[۵۸]. روش تخمین مزبور به صورت زیر تعریف می‌شود:

57 - Trace

$$\hat{\beta}_r = [X'X + KI]^{-1} X'Y$$

K ماتریس قطری است که عنصر روی قطر آن عددی اختیاری است. K طوری انتخاب می‌شود که، نتایج تخمین‌ها پایدار شوند. گرچه $\hat{\beta}_r$ داری تورش است.

چون:

$$E(\hat{\beta}_r) = (X'X + KI)^{-1}(X'X)\beta$$

ولی واریانس آن کوچکتر از واریانس تخمین‌زننده‌های OLS است.

$$VC(\hat{\beta}_r) = \sigma^2 (X'X + KI)^{-1}(X'X)(X'X + KI)^{-1}$$

مثال: برای داده‌های آماری زیرکه میزان اشتغال کل را برروی عرض مبدا، روند GNPD ،GNP و مخارج نیروی نظامی، رگرس نموده است، بیان می‌کند.

	ثابت	T	GNPD	GNP	Armed Forced
K= 0(OLS)	—	-576	-19.76	0.064394	0.010
K=0.0000001	—	-564.21	-25.77	0.064439	-0.0043

بیان دیگر این روش آنست که مقدار ثابت K را به واریانس متغیرهای توضیحی اضافه کنیم، هر چند تخمین پارامترها تورش‌دار است ولی MSE کاهش پیدا می‌کند. یکی از راههای انتخابی K، ضریب لاگرانژ در مدل OLS مقید است، بطوریکه مثلاً:

$$\sum_{i=1}^{k} \beta_i^2 = C$$

مثلاً در یک مدل دو متغیره داریم:

$$Y = \beta_1 X_1 + \beta_2 X_2 + u$$

$$s.t: \beta_1^2 + \beta_2^2 = C$$

$$Min: \Sigma(y - \beta_1 X_1 - \beta_2 X_1)^2 + k(\beta_1^2 + \beta_2^2 - C)$$

$$2\Sigma(y - \beta_1 X_1 - \beta_2 X_1)(-X_1) + 2k\beta_1 = 0$$

$$2\Sigma(y - \beta_1 X_1 - \beta_2 X_1)(-X_2) + 2k\beta_2 = 0$$

$$(\Sigma X_1^2 + K)\beta_1 + \Sigma X_1 X_2 \beta_2 = \Sigma X_1 y$$

$$(\Sigma X_1 X_2)\beta_1 + \Sigma(X_2^2 + K)\beta_2 = \Sigma X_2 y$$

در واقع معادلات فوق، معادله نرمال مربوط به روش برآورد Ridge می‌باشد و K ضریب لاگرانژ است.

راه حل دیگری که برای مشکل هم خطی پیشنهاد می‌شود، روش: Principale Component Regression است. فرض کنید K تا متغیر داریم، می‌توانیم روابط خطی بین این مغیرها را بصورت زیر در نظر بگیریم

.

$$Z_1 = a_1 X_1 + a_2 X_2 + ... + a_k X_k$$

58 -Ridge.

$$Z_2 = b_1 X_1 + b_2 X_2 + ... + b_k X_k$$
$$Z_1 = a_{1\,1} + a_2 X_2 + ... + a_k X_k$$

و بهمین ترتیب $Z_1 = x a_1$

بردار a_1 بردار ریشه ویژه $X'X$ متناظر با ریشه ویژه λ_1 است، که λ_1 بزرگرین ریشه مشخصه است.

اگر کل تغییرات در X با اثر $X'X$ تعریف شود، تغییرات $X'X$ با $W_1 = \frac{\lambda_1}{\Sigma \lambda_1}$ قابل توضیح است.

فرض کنید a_i ها را طوری انتخاب کنیم که واریانس Z_1 ماکزیمم شود با قید اینکه مجموع مجذورات a_i برابر با یک باشد. $a_1^2 + a_2^2 + ... + a_k^2 = 1$ که شرط نرمال کردن[۵۹] است.

آنگاه Z_1 را اولین Principle Component گویند. پس PC تابعی خطی از x هاست که دارای بیشترین واریانس است. جریان ماکزیمم نمودن واریانس توابع خطی Z با قید اینکه مجموع ضرایب x_i ها برابر با یک باشد منجر به K تا جواب خواهد شد، بطوریکه K تا تابع خطی خواهیم داشت که PC های x ها خوانده می‌شوند. پس از آن واریانس‌های Z_i ها را می‌توانیم بصورت زیر مرتب کنیم.

$$Var(Z_1) > Var(Z2) > .... > Var(Z_k)$$

Z_1 با بیشترین واریانس اولینPC خوانده می‌شود. Z_2 با بیشترین واریانس دومین PC خوانده می‌شود و بهمین ترتیب این PC ها دارای ویژگی های زیرند:

۱- $Var(Z_1) + Var(Z_2) + ... + Var(Z_k) = Var(X_1) + Var(X_2) + ... + Var(X_k)$

۲- برخلاف X ها که باهم رابطه دارند، Z ها برهم عمودند یا رابطه ندارند. لذا بعضی پیشنهاد می‌کنند که بجای رگرسیون Y بر رویX ها، Y را بر روی Z ها رگرس کنیم. ولی این روش صحیح نیست چراکه به جوابهای مشابه خواهیم رسید. اما اگرY را روی زیر مجموعه‌ای ازZ ها کنیم، پیشرفتی حاصل می‌شود گرچه دارای مشکلاتی بشرح زیراست:

۱- اگرچه اولین PC بالاترین واریانس را دارد ولی معلوم نیست بیشترین همبستگی را با Y داشته باشد.

۲- ممکن است بگوئیم آن PC هایی را انتخاب می‌کنیم که بیشترین رابطه را باY دارند و بقیه را کنار می‌گذاریم، این کار دربارهX ها هم عملی است.

۳- ترکیب خطیX ها مفهوم اقتصادی ندارد. یعنی مثلاً : (قیمت)۳+(درآمد)۲ مفهومی ندارد.

PC ما را راهنمایی می‌کند که چند منبع تفسیر مستقل درداخل مدل وجود دارد. فرض کنید چندین نرخ بهره داخل مدل وجود دارد، ولی روشPC نشان می‌دهد که دو تا PC ، ۹۹درصد تغییرات نرخ بهره را نشان می‌دهند.می‌گوئیم فقط دو متغیر پنهان وجود دارد که همه تغییرات نرخهای بهره رانشان می‌دهد.

59 - Normalization.

فصل سیزدهم

۱۳-۱ سیستم معادلات همزمان خطی[۶۰]

چنانچه متغیر وابسته (درونزا) یک معادله در سمت راست معادلات دیگر ظاهر شـود، معـادلات همزمـان خواهیم داشت. دو مسئله اساسی در معادلات همزمان وجود دارد که عبارتسـت از تـورش همزمـانی و مسـئله تشخیص. برای آنکه به این دو مشکل در معادلات همزمان پی ببریم به مثال زیر توجه کنیم:

مدل ساده کینزی زیر را در نظر بگیریم:

$$۱)\quad C_t = \alpha Y_t + \varepsilon_t \qquad \varepsilon_t \sim i.i.dN\left(0, \sigma_\varepsilon^2\right)$$

$$۲)\quad Y_t = C_t + I_t \qquad I_t = I_0$$

چنانچه پارامتر α را در معادله (۱) به روش OLS تخمین بزنیم داریم:

$$۳)\quad \hat{\alpha} = \frac{\sum C_t Y_t}{\sum Y_t^2}$$

$$۴)\quad \hat{\alpha} = \alpha + \frac{\sum Y_t \varepsilon_t}{\sum Y_t^2}$$

---

60 . Linear Simultaneous Equations Model

تورش‌دار بودن یا نبودن $\hat{\alpha}$ بستگی به امید ریاضی حاصلضرب ε_t و Y_t دارد. اگر $E(Y_t\varepsilon_t)$ برابر صفر باشد، $\hat{\alpha}_{OLS}$ بدون تورش و چنانچه $E(Y_t\varepsilon_t)$ مخالف صفر باشد، $\hat{\alpha}_{OLS}$ دارای تورش است. برای آنکه به این مسئله پی ببریم، کافیست آن را محاسبه کنیم.

$$Y_t = \alpha Y_t + I_0 + \varepsilon_t$$

$$Y_t = \frac{I_0}{1-\alpha} + \frac{\varepsilon}{1-\alpha}$$

$$Y_t\varepsilon_t = \frac{I_0\varepsilon_t}{1-\alpha} + \frac{\varepsilon_t^2}{1-\alpha}$$

$$E(Y_t\varepsilon_t) = 0 + \frac{\sigma_\varepsilon^2}{1+\alpha} \neq 0$$

ملاحظه می‌کنیم $\hat{\alpha}_{OLS}$، دارای تورش است. به این تورش، تورش همزمانی[61] می‌گویند.

لذا با جایگزاری در رابطه (۴) داریم:

$$\hat{\alpha} = \alpha + \frac{(1-\alpha)}{1+\frac{I_0^2}{\sigma_\varepsilon^2}}$$

سئوال این است که آیا $\hat{\alpha}$ سازگار است؟

$$P\lim\hat{\alpha} = \alpha + \frac{P\lim\frac{1}{T}\sum Y_t}{P\lim\frac{1}{T}\sum Y_t^2}$$

صورت جمله دوم به سمت $\frac{\sigma_\varepsilon^2}{1-\alpha}$ و مخرج به سمت $\frac{I_0^2+\sigma_\varepsilon^2}{(1-\alpha)^2}$ میل می‌کند.

لذا داریم: $P\lim\hat{\alpha} = \alpha + \frac{(1-\alpha)}{1+\frac{I_0^2}{\sigma_\varepsilon^2}}$

نتیجه می‌گیریم که $\hat{\alpha}_{OLS}$ در معادلات همزمان دارای تورش و ناسازگار است. برای آنکه به یک تخمین زننده سازگار برسیم میتوانیم روش متغیر ابزاری را مورد استفاده قرار دهیم.

61 . Simultaneous bias

۱۳-۲ روش متغیرهای ابزاری :

مدل $y_t = \beta x_t + u_t$ را در نظر بگیریم، همانطور که در فروض رگرسیون آمده است درصورتیکه $Cov(x_t, u_t) \neq 0$ باشد، تخمین‌زننده‌های OLS ناسازگار خواهند بود. متغیرهای ابزاری در اینجا عبارتست از متغیری که با u_t ارتباط نداشته باشد ولی با متغیر x_t دارای همبستگی بالایی باشد. فرض کنیم z_t چنین متغیری باشد. بطوریکه $Cov(z_t, u_t) = 0$ باشد، در آنصورت داریم:

$$\sum z_t y_t = \beta \sum z_t x_t + \sum z_t u_t$$

چون $P\lim \frac{\sum z_t u_t}{T} \to 0$ لذا داریم:

$$\hat{\beta}_{IV} = \frac{\sum z_t y_t}{\sum z_t x_t}$$

حال مدل را بصورت چند متغیره در نظر بگیریم:

$$Y = X\beta + U$$

همانطور که می‌دانیم: $P\lim \frac{X'U}{T} \to 0$ و طبق فرض $P\lim \frac{X'X}{T} \to Q$

لذا تخمین‌زننده $\hat{\beta}_{OLS}$ نیز یک تخمین‌زننده IV است که متغیر ابزاری رگرسورهای X است چون :

$$P\lim \hat{\beta} = \beta$$

حال اگر در مدلی $P\lim \frac{X'U}{T}$ به سمت صفر میل نکند، مثلاً یکی از عناصر X، Y_{t-1} باشد و در نتیجه بین Y_{t-1} و U_t رابطه برقرار باشد، باید متغیر یا متغیرهایی در قالب ماتریس $W_{T\times K}$ پیدا کنیم که با جملات اخلال رابطه‌ای نداشته باشند و در عین حال با X همبستگی داشته باشد.

یعنی : $E(W_{ti}U_t) = 0$

در این صورت: $W'Y = W'X\beta + W'U$

اگر $P\lim \frac{W'U}{T} \to 0$ باشد. در آن صورت $\hat{\beta}_{IV}$ بصورت زیر تعریف می‌شود.

$$\hat{\beta}_{IV} = (W'X)^{-1} W'Y$$

پس داریم:

$$\hat{\beta}_{IV} = \beta + (W'X)^{-1} W'U$$

$$P\lim \hat{\beta}_{IV} = \beta + P\lim \left(\frac{W'X}{T}\right)^{-1} P\lim \frac{W'U}{T}$$

$$P\lim \hat{\beta}_{IV} = \beta + Q_{wx}^{-1}.0 = \beta$$

مثال: مجددا تابع درآمد و مصرف مدل کینز را که برای سادگی عرض از مبدا آنرا کنار گذاشته ایم در نظر بگیریم:

$$C_t = \beta Y_t + \varepsilon_t$$
$$Y_t = C_t + I_t$$

چنانچه $\hat{\beta}$, از طریق OLS تخمین زده شود، دارای تورش خواهد بود و ناسازگاراست.
چنانچه متغیر I_t را بعنوان متغیر ابزاری در نظر بگیریم داریم:

$$\hat{\beta}_{IV} = \frac{\sum I_t C_t}{\sum I_t Y_t} = \beta + \frac{\sum I_t \varepsilon_t}{\sum I_t Y_t}$$

$$P\lim \hat{\beta}_{IV} = \beta + \frac{P\lim \frac{\sum I_t \varepsilon_t}{T}}{P\lim \frac{\sum I_t Y_t}{T}}$$

$$P\lim \frac{\sum I_t \varepsilon_t}{T} \to 0$$

چون:

$$P\lim \frac{\sum I_t Y_t}{T} \to Q_{Iy}$$

ثابت

$$P\lim \hat{\beta}_{IV} = \beta$$

در

نتیجه:

۱۳-۳ مسئله تشخیص[۶۲]:

در مدل $y = X\beta + u$ اگر X ها غیرتصادفی باشند و $u \sim N(0, \sigma^2 I)$ اگر ما توزیع y را داشته باشیم و X ها هم معلوم باشند، قادر خواهیم بود مقادیر واقعی β را پیدا کنیم، بدین ترتیب که چون :

$$Y \sim N(X\beta, \sigma^2 I)$$

حال اگر $X\beta$ را داشته باشیم از رابطه زیر می‌توانیم β را پیدا کنیم.

$$Ey = X\beta$$

سئوال این است که آیا رابطه فوق برای β بطور منحصر بفردی قابل حل است؟

اگر پاسخ مثبت باشد، β قابل تشخیص و شناسایی است.

چنانچه پاسخ منفی است، β قابل تشخیص نیست.

در معادلات غیرهمزمان می‌گفتیم اگر Ey معلوم و $X\beta$ معلوم، چنانچه دچار هم‌خطی نباشیم قادریم β را تشخیص بدهیم. شرطی که برای این کار قرار می‌دادیم این بود که $Rank(X) = k$ یعنی ماتریس X دارای K ستون مستقل از هم باشد. پس چون این شرط را قرار می‌دادیم، مدل خطی غیرهمزمان قابل تشخیص بود.

62 . Indentification

۱۳-۴ تشخیص پارامترهای فرم ساختاری:

مدل عرضه و تقاضای زیر را در نظر بگیریم:

مدل ۱ –

تابع تقاضا $q_t^D = a_1 + b_1 p_t + c_1 l_t + u_{t1}$

تابع عرضه $q_t^S = a_2 + b_2 p_t + c_2 r_t + u_{t2}$

$$q_t = q_t^S \equiv q_t^D$$

l_t و r_t به ترتیب عبارتند از میزان درآمد و بارندگی که هر دو متغیرهای برونـزا هســتند، در ایـن مـدل قیمت و و مقدار متغیرهای درونزا هستند.

$$U_{t0} = [u_{t1}, u_{t2}], i.i.d \sim N(0, \Sigma)$$

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}$$

با استفاده از شرط تعادل داریم:

$$\begin{cases} q_t^p = a_1 + b_1 p_t + c_1 l_t + u_{t1} \\ q_t^s = a_2 + b_2 p_t + c_2 r_t + u_{t2} \end{cases}$$

فرم بالا را فرم ساختاری[۶۳] می‌نامند.

پارامترهای Σ, c, b, a را پارامترهای مدل ساختاری می‌گویند.

با جایگزاری و حل مدل یا فرم بالا می‌توانیم به فرم حل شده[۶۴] دست یابیم:

$$\begin{cases} q_t = \pi_1 + \pi_2 l_t + \pi_3 r_t + v_{t1} \\ q_t = \pi_4 + \pi_5 l_t + \pi_6 r_t + v_{t2} \end{cases}$$

$$V_{t0} = [v_{t1}, v_{t2}] i.i.d \sim N(0, \Omega)$$

$$\Omega = \begin{bmatrix} \omega_{11} & \omega_{12} \\ \omega_{21} & \omega_{22} \end{bmatrix}$$

بطوریکه:

$$\pi_1 = \frac{a_1 b_2 - a_2 b_1}{b_2 - b_1}, \pi_2 = \frac{c_1 b_2}{b_2 - b_1}, \pi_3 = \frac{-c_2 b_1}{b_2 - b_1}$$

$$\pi_4 = \frac{a_1 - a_2}{b_2 - b_1}, \pi_5 = \frac{c_1}{b_2 - b_1}, \pi_6 = \frac{-c_2}{b_2 - b_1}$$

$$V_{t1} = \frac{1}{b_2 - b_1}(u_{t1} b_2 - u_{t2} b_1), V_{t2} = \frac{1}{b_2 - b_1}$$

63 . Structural Form
64 . Reduced Form

$$\sigma_{11}=\omega_{11}+\omega_{12}+\omega_{21}+\omega_{22}$$
$$\sigma_{12}=-b_1(\omega_{11}+\omega_{21})-b(\omega_{12}+\omega_{22})$$
$$\sigma_{22}=b_1^2\omega_{11}+2b_1b_2\omega_{12}+b_2^2\omega_{22}$$

پارامترهای Ω,π را پارامترهای فرم حل شده می‌نامند.

فرض کنید r_t, e_t غیرتصادفی باشند و توزیع q_t, p_t معلوم باشد.

الف: حال مسئله شناسائی در فرم حل شده بصورت زیر قابل بررسی است. فرض کنید از توزیع متغیرهای درونزا اطلاع داشته باشیم (حداقل دو گشتاور اول) $y_{t0}=[q_t, p_t]$

آیا می‌توانیم بطور منحصر بفردی مقادیر $\pi_1,\pi_2,\pi_3,\pi_4,\pi_6$ و $\omega_{11},\omega_{12},\omega_{22}$ را از فرم حل شده بدست آوریم؟ به عبارت دیگر، اگر دو سری پارامتر بصورت زیر داشته باشیم:

$$\begin{cases}\pi_1^{(1)},\pi_2^{(1)},\dots\pi_6^{(1)},\omega_{11}^{(1)},\dots\omega_{22}^{(1)}\\ \pi_1^{(2)},\pi_2^{(2)}\dots\pi_6^{(2)},\omega_{11}^{(2)}\dots\omega_{22}^{(2)}\end{cases}$$

که هر دو سری توزیع مشابهی را برای y_{t0} نشان می‌دهند.

آنگاه مشاهده y_{t0} و متغیرهای برونزا $Z_{t0}=[1,r_t,1_t]$ به ما اجازه نخواهد داد تا ساختار موردنظر را تعیین کنیم، حتی اگر مشاهدات نامحدود باشد. حال:

$$\mu=E=\begin{bmatrix}y_{10}\\y_{20}\\\vdots\\y_{T0}\end{bmatrix}=\begin{bmatrix}z_{10}\\z_{20}\\\vdots\\z_{T0}\end{bmatrix}\begin{bmatrix}\pi_1&\pi_4\\\pi_2&\pi_5\\\pi_3&\pi_6\end{bmatrix}$$

اگر مرتبه ماتریس Z برابر ۳ باشد یعنی کامل، آنگاه با مشاهده $\mu=E(y)$، می‌توانیم π را محاسبه کنیم و بطور منحصر بفرد بدست بیاوریم.

ماتریس واریانس- کوواریانس y_{t0} برابر است با واریانس کوواریانس U_{t0} . لذا می‌توانیم ماتریس واریانس- کوواریانس V_{t0} یعنی Ω را نیز بدست آوریم. در نتیجه اگر ماتریس مشاهدات بر روی متغیرهای برونزا دارای مرتبه کامل باشد و جمله اخلال $i.i.d$ باشند، پارامترهای فرم ساختاری قابل شناسایی هستند.

ب- مسئله شناسایی در فرم ساختاری: گیریم توزیع متغیرهای درونزا (حداقل دو گشتاور اول) را می‌دانیم، یعنی توزیع $y_{t0}=[q_t, p_t]$ معلوم است. آیا می‌توانیم بطور منحصر بفرد و یک بیک به پارامترهای فرم ساختاری Σ,c,b,a برگردیم؟

مجدداً این بحث مطرح می‌شود که اگر دو سری پارامتر برای فرم ساختاری وجود دارد که توزیع مشابهی را نشان می‌دهد، ما نمی‌توانیم بین دو ساختار تفاوت قائل شویم و ساختار موردنظر را تعیین کنیم، حتی اگر حجم نمونه نامحدود باشد.

همانطور که قبلاً نشان داده شد، داشتن دو گشتاور اولیه از توزیع متغیرهای وابسته (q_t, p_t در اینجا) براساس فروض مدل ساختاری، برابر است با: π,Ω

لذا ، سئوال درباره شناسایی پارامترهای فرم ساختاری به این سئوال تبدیل می‌شود که اگر، $\pi_1,\dots,\pi_6$ و $\omega_{11},\dots,\omega_{22}$ را داریم، آیا می‌توانیم بطور منحصر بفرد و یک به یک مقادیر پارامترهای فرم ساختاری یعنی $\sigma_{11},\sigma_{12},\sigma_{22},a_1,b_1c_1,a_2,b_2,c_2$ را بدست آوریم؟

توجه داریم که اگر b_2, b_1 و $\omega_{11}, \omega_{12}, \omega_{22}$ را بدست آورم، مقادیر $\sigma_{11}, \sigma_{12}, \sigma_{22}$ نیز برای ما معلوم می‌شود. در نتیجه سئوال منجر به این سئوال می‌شود که آیا $\pi_1, \ldots, \pi_6$ نشانگر مقادیر منحصر بفرد برای $a_1, \ldots, c_2$ می‌باشد؟

برای مدل فوق داریم:

$$b_1 = \frac{\pi_3}{\pi_6} \quad , \quad b_2 = \frac{\pi_2}{\pi_5} \quad , \quad c_2 = -\pi_6(b_2 - b_1)$$
$$c_1 = -\pi_5(b_1 - b_2) \quad , \quad a_1 = \pi_1 - b_1\pi_4 \quad , \quad a_2 = \pi_1 - b_2\pi_4$$
$$b_2 \neq b_1 \quad , \quad c_2 \neq 0$$

و چون Ω معلوم است، δ_{ij} ، بطور منحصر بفرد، تعیین می‌شوند.

لذا پارامترهای فرم ساختاری مدل (۱) قابل شناسایی هستند.

حال مدل دیگری از عرضه و تقاضا را در نظر بگیریم:

مدل ۲-

$$q^D = a_1 + b_1 p_t + c_1 l_t + u_{t1}$$
$$q_t^s = a_2 + b_2 p_t + \qquad u_{t2}$$
$$q_t^D = q_t^s = q_t$$

فرم خلاصه شده بصورت زیر است:

$$q_t = \pi_1 + \pi_2 l_t + v_{t1} \qquad \pi_1 = \frac{a_1 b_2 - a_2 b_1}{b_2 - b_1} \quad , \quad \pi_2 = \frac{c_1 b_2}{b_2 - b_1}$$
$$p_t = \pi_4 + \pi_5 l_t + v_{t2} \qquad \pi_4 = \frac{a_1 - a_2}{b_2 - b_1} \quad , \quad \pi_5 = \frac{c_1}{b_2 - b_1}$$

که مانند مدل (۱) محاسبه شده است، ملاحظه می‌شود تفاوت فرم خلاصه شده فوق با فرم خلاصه شده مدل (۱) تنها در حذف متغیر r_t در معادله عرضه می‌باشد. پارامترهای فرم خلاصه شده باز هم قابل شناسایی هستند. اما برای رسیدن به پارامترهای فرم ساختاری فقط ۴ پارامتر معلوم در فرم حل شده داریم در حالیکه در مدل ساختاری ۵ پارامتر دیده می‌شود $(a_1, b_1, c_1, a_2, b_2)$ آشکار است که سیستم کم‌تشخیص است[۶۵]. اما علیرغم آن پارامترهای فرم ساختاری تابع عرضه قابل شناسایی هستند. یعنی:

$$b_2 = \frac{\pi_2}{\pi_5} \quad , \quad a_2 = \pi_1 - b_2\pi_4$$

لذا تابع عرضه قابل تشخیص ولی تابع تقاضا قابل تشخیص نیست.

اما اگر فرض کنیم $b_1 = -b_2$ یا به عبارت دیگر مجموع کشش عرضه و تقاضا برای کالای خاص قرینه است در آنصورت داریم:

$$c_1 = \pi_5 - (b_2 - b_1), a_1 = \pi_1 - b_2\pi_4$$

65 . Underidentified

مدل ۳-

و بالاخره تابع عرضه و تقاضای زیر را در نظر بگیریم:

$$q_t^D = a_1 + b_1 p_t + c_1 l_t + d_1 r_t + u_{t1}$$
$$q_t^S = a_2 + b_2 p_t + u_{t2}$$
$$q_t^D = q_t^S = q_t$$

فرم خلاصه شده مدل بصورت زیر است:

$$q_t = \pi_1 + \pi_2 l_t + \pi_3 r_t + v_{t1}$$
$$p_t = \pi_4 + \pi_5 l_t + \pi_6 r_t + v_{t2}$$
$$\pi_1 = \frac{a_1 b_2 - a_2 b_1}{b_2 - b_1}, \pi_2 = \frac{c_1 b_2}{b_2 - b_1}, \pi_3 = \frac{d_1 b_2}{b_2 - b_1}$$
$$\pi_4 = \frac{a_1 b_2}{b_2 - b_1}, \pi_5 = \frac{c_1}{b_2 - b_1}, \pi_6 = \frac{d_1}{b_2 - b_1}$$

در این مدل، دو فرم برای b_2 و a_2 داریم یعنی:

$$b_2 = \frac{\pi_3}{\pi_6}$$
یا $b_2 = \pi_2 / \pi_5$
$$a_2 = \pi_1 - b_2 \pi_4$$

و برای d_1, c_1, b_1, a_1 هیچ جوابی نخواهیم یافت.

در اینجا تابع عرضه بیش از حد قابل شناسایی است ولی تابع تقاضا قابل شناسایی نیست.

پس بین سه حالت دقیقاً قابل تشخیص[66] ، فوق تشخیص[67] و کم تشخیص[68] تفاوت قائل هستیم.

ملاحظه کنید، فرمولهای فوق را برای بدست آوردن و تخمین پارامترهای فرم ساختاری استفاده می‌کنیم اما در مدل (۳) داریم:

$$\hat{b}_2 = \frac{\hat{\pi}}{\hat{\pi}}, \hat{\hat{b}}_2 = \frac{\hat{\pi}_2}{\hat{\pi}_5}$$

و لذا $\hat{b}_2 \neq \hat{\hat{b}}_2$ یعنی π ها جواب‌های منحصر بفردی بدست نمی‌دهند. در اینصورت می‌گوئیم پارامترها فوق تشخیص هستند.

۱۳-۵ نمادها ، فروض و قضایای حد مرکزی:

چنانچه مشاهده t ام از معادله i ام را در نظر بگیریم:

معادله i ام برای T مشاهده عبارتست از :

---

66 . Exact Identified
67 . Over Identified
68 . Under Identified

$$y_{ti}=\sum_{j=1}^{G}y_{tj}b_{ji}+\sum_{j=1}^{K}Z_{tj}C_{ji}+U_{ti} \qquad t=1,\dots,T$$

بطوریکه: $b_{ii=0}$

معادله i ام برای T مشاهده عبارتست از :

$$y_{.i}=Yb_{.i}+ZC_{.i}+U_{.i}$$

بطوریکه: $\underset{T\times G}{Y}=\left[y_{.1},\dots,y_{.G}\right] \quad ,\underset{T\times K}{Z}=\left[z_{.1},\dots,z_{.k}\right]$

G_i تعداد متغیرهای درونزا و K_i تعداد متغیرهای برونزا در معادله i ام است.

Y_i عبارتست از بردار متغیرهای درونزا که در معادله i ام ظاهر شده‌اند.

Z_i عبارتست از بردار متغیرهای از پیش تعیین شده که در معادله i ام ظاهر شده‌اند.

β_i و γ_i را بصورت زیر تعریف می‌کنیم:

$$\beta_i : G_i\times 1$$
$$\gamma_i : K_i\times 1$$

معادله i ام را می‌توانیم بصورت زیر بنویسیم:

$$y_{.i}=Y_i\beta_i+Z_i\gamma_i+u_{.i}$$

$$y_{.i}=X_i\delta_i$$

بطوریکه: $X_i=\left[Y_i,Z_t\right],\quad \delta_i=\begin{bmatrix}\beta_i\\ \gamma_i\end{bmatrix}$

حال فرض کنیم بخواهیم ستون j ام را از ماتریس Y انتخاب کنیم، خواهیم داشت:

$$y_{.i}=Y.e_{.j} \qquad e_{.j}=\begin{bmatrix}0\\0\\1\\0\\ \vdots\\0\end{bmatrix}\leftarrow j$$

$e_{.j}$ یک بردار $G\times 1$ است که همه عناصر آن بجز عنصر j ام صفر است و عنصر j ام برابر یک است.

$$\underset{T\times G_i}{Y_i}=\left[y_{.i},\dots,y_{.i}G_i\right]=Y\left[e_{.i1},\dots,e_{.i}G_1\right]=\underset{T\times K}{Y}\underset{G\times G_i}{L_{i1}}$$

و بهمین ترتیب

$$\underset{T\times K_i}{Z_i}=\underset{T\times K}{Z}\underset{K\times K_i}{L_{i2}}$$

L_{i1} و L_{i2} را ماتریس‌های انتخاب‌گر می‌نامیم.

برای مثال:

$$b_{.i} = L_{i1}\beta_i$$
$$c_{.i} = L_{i2}\gamma_i$$

لذا فرم ساختاری مدل را می‌توانیم بصورت زیر بنویسیم:

$$[y_{.1},\dots,y_{.G}] = Y[b_{.1},\dots,b_{.G}] + Z[c_{.1},\dots,c_{.G}] + [u_{.1},\dots,u_{.G}]$$

اگر ماتریسهای Y و U را برداری نمائیم داریم:

$$\begin{bmatrix} y_{.1} \\ y_{.2} \\ \vdots \\ y_{.G} \end{bmatrix} = \begin{bmatrix} x_1 & & 0 \\ & \ddots & \\ 0 & & x_G \end{bmatrix} \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_G \end{bmatrix} + \begin{bmatrix} u_{.1} \\ u_{.2} \\ \vdots \\ u_{.G} \end{bmatrix}$$

یا:

$$y = X\delta + u$$

بطوریکه:

$$y = Vec(y), \quad u = Vec(U)$$

$$Vec\begin{pmatrix} B \\ C \end{pmatrix} = \begin{bmatrix} L_1 & & 0 \\ & \ddots & \\ 0 & & L_G \end{bmatrix}\delta$$

$$Vec\begin{pmatrix} B \\ C \end{pmatrix} = \begin{bmatrix} L_{i1} & 0 \\ 0 & L_{i2} \end{bmatrix}$$

δ: بردار پارامترها بدون اعمال قید است.

$$\underset{T\times G}{Y} = \underset{T\times G}{Y}\,\underset{G\times G}{B} + \underset{T\times G}{Z}\,\underset{G\times G}{C} + \underset{T\times G}{U}$$

حال چنانچه Yرا به سمت چپ برده و از آن فاکتور بگیریم داریم:

$$Y(I-B) = ZC + U$$
$$Y = ZC(I-B)^{-1} + U(I-B)^{-1}$$

فرم مقابل را فرم خلاصه شده می‌گویند.

$$Y = Z\Pi + V$$

بطوریکه:

$$\Pi = C(I-B)^{-1}, \quad V = U(I-B)^{-1}$$

۱۳- ۶ فروض مدل :

۱- بردار $u'_{t.}$ ؛ $i.i.d$ هستند؛

یعنی جملات اخلال در طول زمان از یکدیگر مستقلند ولی بر روی معادلات خیر.

$$E(u'_{t.}u_{s.}) = 0 \qquad \text{for} \quad t \neq s$$

$$E(u'u) = E\left\{\sum_{t=1}^{T} u'_{t.}u_{t.}\right\} = T.\Sigma$$

$$E(u_{.i}u'_{.j}) = \sigma_{ij} I_T$$

$$E(uu') = E\begin{bmatrix} u_{.1} \\ u_{.2} \\ \vdots \\ u_{.G} \end{bmatrix} [u'_{.1}, \dots, u'_{.G}] = \Sigma \otimes I_T$$

۲- ماتریس $I - B$ غیرمنفرد است.

۳- فرض کنید W زیرماتریسی از Z باشد، متناظر با متغیرهای برونزا (فقط برونزا) آنگاه:

غیرمنفرد و محدود $$Q^* = \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} w'_{t.} w_{t.} = \lim_{T\to\infty} \frac{1}{T} w'w$$

محدود برای $$Q^{**} = \lim \frac{1}{T}\sum_{t=1}^{T-K} w'_{t.} w_{t+K} k = 1,2,\dots$$

۴- شرایط پایداری:

$$\mathrm{Y(I-B) - ZC - U = Y(I-B) - WC_0 - Y_{-1}C_1 - \dots - Y_{-H}C_H - U = 0}$$

بطوریکه Y_{-i} عبارتست از مشاهده روی متغیر تأخیری Y برای دوره i ام و H ماکزیمم تاخیراست. آنگاه فرض می‌کنیم که قدر مطلق ریشه معادله دترمینان

$$\left|\lambda^H (I-B) - \lambda^{H-1} C_1 \dots C_H\right| = 0$$

کوچکتر از یک باشد.

۵- عناصر ماتریس W بطور یکنواخت *Bounded* ؛ $|w_{ti}| \le c < \infty$ باشند.

فروض مربوط به قضایای حدی مرکزی:

Q محدود و غیرمنفرد است. بطوریکه Q^{-1} وجود دارد.

$$Q = P\lim \frac{Z'Z}{T}$$

$$Vec\left(\frac{Z'u}{\sqrt{T}}\right) = \begin{bmatrix} \frac{Z'u_{.1}}{\sqrt{T}} \\ \vdots \\ \frac{Z'u_{.G}}{\sqrt{T}} \end{bmatrix} \xrightarrow{in.d} N(0, \Sigma \otimes Q)$$

اثبات: Schonfild(1971)

$$\xi = \mathrm{Vec}\left(\frac{Z'u}{\sqrt{T}}\right) = \left(I \otimes \frac{Z'}{\sqrt{T}}\right)u$$

$$E(\xi) = 0$$

$$E(\xi\xi') = E(I \otimes \frac{Z'}{\sqrt{T}})uu'(I \otimes \frac{Z}{\sqrt{T}})$$

$$= \left(I \otimes \frac{Z'}{\sqrt{T}} \right) (\Sigma \otimes I) \left(I \otimes \frac{Z}{\sqrt{T}} \right) = \Sigma \otimes \frac{Z'Z}{T}$$

نتیجه ۱-

$$\frac{Z'u_{.i}}{\sqrt{T}} \to N(0, \sigma_{ii} Q)$$

نتیجه ۲-

$$P\lim \frac{Z'u_{.i}}{\sqrt{T}} = 0$$

$$P\lim \frac{Z'U}{T} = 0$$

نتیجه ۳-

$$P\lim \frac{Z'v_{.i}}{T} = 0$$

$$P\lim \frac{Z'V}{T} = 0$$

اثبات:

$$P\lim \frac{Z'V}{T} = P\lim \frac{Z'U}{T} (I - B)^{-1} = 0$$

بر اساس نمادهای اشاره شده و فروضی که بیان گردید و قبل از ورود به بحث تخمین لازم است موضوع شناسایی پارامترها بطور کلی و شرط لازم و کافی مورد بررسی قرار گیرد.

تشخیص در معادله i ام :

توجه ۱- اگر هیچگونه محدودیتی روی مدل ساختاری اعمال نگردد، مدل قابل تشخیص نیست. تنها با اعمال محدودیت می‌توانیم پارامترهای بعضی یا همه معادلات را قابل تشخیص کنیم.

محدودیتهایی مانند : $b_{ji} = 0$ و $C_{ji} = 0$ یعنی متغیر درونزای j ام و متغیر برونزای j ام در معادله i ام ظاهر نشوند.

برای سادگی معادله یکم را در نظر بگیریم:

$$y_{.1} = Y_1 \beta_1 + Z_1 \gamma_1 + u_{.1}$$

Z_1 ماتریس مشاهدات بر روی متغیرهای برونزا ظاهر شده در معادله یکم است.

$$Z_1 = [Z_{.1}, \ldots, Z_{.K_1}]$$

Y_1 ماتریس مشاهدات بر روی متغیرهای درونزا ظاهر شده در معادله یکم است.

$$Y_1 = [Y_{.2}, \ldots, y_{.G}]$$

$$Y = [y_{.1}, Y_1, Y_1^*]$$

$$Z = [Z_1, Z_1^*]$$

متغیرها یا پارامترهایی که با علامت ستاره مشخص شده اند یعنی اینکه در سیستم معـادلات وجـود دارند ولی در معادله مورد نظر ظاهر نشده اند.

فرم خلاصه شده عبارتست از :

$$Y = Z\Pi + V$$

$$\Pi = C(I-B)$$

$$\Pi(I-B) = C$$

$$\Pi(e_{.1} - b_{.1}) = C_{.1}$$

فرم خلاصه شده :

$$[y_{.1}, Y_1, Y_1^*] = [Z_1, Z_1^*]\begin{bmatrix} \underset{K_1\times 1}{\Pi_{11}} & \underset{K_1\times G_1}{\Pi_{12}} & \underset{K_1\times G-G_1-1}{\Pi_{13}} \\ \underset{K-K_1}{\Pi_{21}} & \underset{(K-K_1)G_1}{\Pi_{22}} & \underset{(K-K_1)(G-G_1-1)}{\Pi_{23}} \end{bmatrix} + [v_{.1}, V, V_1^*]$$

معادله فرم خلاصه شده مربوط به متغیرهای درونزای معادله یکم عبارتست از:

$$Y_1 = Z_1\Pi_{12} + Z_1^*\Pi_{22} + V_1$$

برای آنکه پارامترهای فرم خلاصه شده معادلات مربوط به معادله یکم را انتخاب کنیم، بصـورت زیـر عمل می‌کنیم:

$$e_{.1} - b_{.1} = \begin{bmatrix} 1 \\ -\beta_1 \\ 0 \end{bmatrix} \begin{matrix} 1\times 1 \\ G_1 \times 1 \\ (G-G)\times 1 \end{matrix}$$

$$C_{.1} = \begin{bmatrix} \gamma_1 \\ 0 \end{bmatrix} \begin{matrix} K_1 \times 1 \\ (K\times K)\times 1 \end{matrix}$$

دستگاه زیر بیانگر رابطه بین پارامترهای فرم ساختاری و فرم خلاصه شده معادله یکم است.:

$$\begin{cases} \Pi_{11} - \Pi_{12}\beta_1 + 0 = \gamma_1 \\ \Pi_{21} - \Pi_{22}\beta_1 + 0 = 0 \end{cases}$$

$$\begin{cases} \Pi_{12}\beta_1 + \gamma_1 = \Pi_{11} \\ \Pi_{22}\beta_1 = \Pi_{21} \end{cases} \qquad \text{(a)}$$

برای آنکه بتوانیم از پارامترهای فرم خلاصه شـده پارامترهـای فـرم سـاختاری را بدسـت آوریـم، در دستگاه فوق ابتدا باید β_1 را بدست می‌آوریم.

$$\beta_1 = \Pi_{21}\Pi_{22}^{-1}$$

پس لازم است Π_{22} دارای مرتبه کامل باشد، اما Π_{22} چیست؟

Π_{22} عبارتست از پارامترهایی که در فرم خلاصه شده وجود دارد و مربوط می‌شوند به ضرایب متغیرهای برونزایی که از معادله اول حذف شده‌اند.

الف – شرط رتبه‌ای[69] :

شرط لازم و کافی برای تشخیص معادله اول این است که مرتبه π_{22} برابر با G_1 باشد.

$$Rank(\Pi_{22}) = G_1$$

اثبات: رابطه (a) بطور منحصر به فرد برای β_1 قابل حل است، اگر و فقط اگر π_{22} رنک کامل ستونی داشته باشد. اگر چنین شد داریم:

$$\gamma_1 = \Pi_{11} - \Pi_{12}\beta_1$$

که دارای جواب منحصر به فرد است.

ب – شرط درجه‌ای[70] :

شرط لازم برای تشخیص معادله اول عبارت از اینکه :

$$K - K_1 \geq G_1 \qquad \text{یا} \qquad K \geq K_1 + G_1$$

یعنی حداقل به تعداد متغیرهای برونزایی که از معادله اول کنار گذاشته شده‌اند متغیر درونزا در آن ظاهر شده باشد یا تعداد متغیرهای برونزای درون سیستم بیشتر از متغیرهای توضیحی معادله می‌باشند.

اثبات: دستگاه دارای K معادله است.

$$\Pi_{22}\beta_1 = \Pi_{12}$$

لذا، به تعداد مجهولات معادله نیاز داریم. یعنی اگر $K = K_1 + G_1$ باشد. معادله دقیقاً قابل تشخیص است و اگر $K > K_1 + G_1$ باشد معادله فوق تشخیص خواهد بود.

مثال:

فرم ساختاری زیر را درنظر بگیریم:

$$\begin{bmatrix} y_{t1} & y_{t2} & \dots & y_{tG} \end{bmatrix} \begin{bmatrix} \gamma_{11} & \gamma_{21} & \cdots & \gamma_{G1} \\ \gamma_{12} & \gamma_{22} & \cdots & \gamma_{G2} \\ \vdots & \vdots & \vdots & \vdots \\ \gamma_{1G} & \gamma_{2G} & \cdots & \gamma_{GG} \end{bmatrix} + \begin{bmatrix} x_{t1} & x_{t2} & \dots & x_{tK} \end{bmatrix} \begin{bmatrix} \beta_{11} & \beta_{21} & \cdots & \beta_{K1} \\ \beta_{12} & \beta_{22} & \cdots & \beta_{K2} \\ \vdots & \vdots & \vdots & \vdots \\ \beta_{1K} & \beta_{2K} & \beta_{34} & \beta_{KG} \end{bmatrix} = \begin{bmatrix} \varepsilon_{t1} & \varepsilon_{t2} & \dots & \varepsilon_{tG} \end{bmatrix}$$

به صورت فرم بسته ماتریسی داریم:

$$Y\Gamma + XB = E$$

بطوریکه: Y ماتریس متغیرهای درونزا T.G و X ماتریس متغیرهای ازپش تعیین شده T.K و Γ ماتریس G.G و B ماتریس K.G است. E ماتریس جملات اخلال و T.G میباشد.

چنانچه مدلی را درنظر بگیریم که دارای ۴ متغیر درونزا و ۴ متغیر ازپیش تعیین شده باشد, بدون اعمال هیچگونه شرطی , سیستم بصورت زیر خواهد بود.

69 . Rank Condition
70 . Order Condition

$$\gamma_{11}y_{t1}+\gamma_{12}y_{t2}+\gamma_{13}y_{t3}+\gamma_{14}y_{t4}+\beta_{11}x_{t1}+\beta_{12}x_{t2}+\beta_{13}x_{t3}+\beta_{14}x_{t4}=\varepsilon_{t1}$$
$$\gamma_{21}y_{t1}+\gamma_{22}y_{t2}+\gamma_{23}y_{t3}+\gamma_{24}y_{t4}+\beta_{21}x_{t1}+\beta_{22}x_{t2}+\beta_{23}x_{t3}+\beta_{24}x_{t4}=\varepsilon_{t2}$$
$$\gamma_{31}y_{t1}+\gamma_{32}y_{t2}+\gamma_{33}y_{t3}+\gamma_{34}y_{t4}+\beta_{31}x_{t1}+\beta_{32}x_{t2}+\beta_{33}x_{t3}+\beta_{34}x_{t4}=\varepsilon_{t3}$$
$$\gamma_{41}y_{t1}+\gamma_{42}y_{t2}+\gamma_{43}y_{t3}+\gamma_{44}y_{t4}+\beta_{41}x_{t1}+\beta_{42}x_{t2}+\beta_{43}x_{t3}+\beta_{44}x_{t4}=\varepsilon_{t4}$$

چنانچه بخواهیم مدل فوق را بر حسب متغیرهای درونزا بصورت صریح دربیاوریم , با قرار دادن γ_{ii} برابر منهای یک آنرا نرمالیزه میکنیم لذا داریم:

$$y_{t1}=\gamma_{12}y_{t2}+\gamma_{13}y_{t3}+\gamma_{14}y_{t4}+\beta_{11}x_{t1}+\beta_{12}x_{t2}+\beta_{13}x_{t3}+\beta_{14}x_{t4}+\varepsilon_{t1}$$
$$y_{t2}=\gamma_{21}y_{t1}+\gamma_{23}y_{t3}+\gamma_{24}y_{t4}+\beta_{21}x_{t1}+\beta_{22}x_{t2}+\beta_{23}x_{t3}+\beta_{24}x_{t4}+\varepsilon_{t2}$$
$$y_{t3}=\gamma_{31}y_{t1}+\gamma_{32}y_{t2}+\gamma_{34}y_{t4}+\beta_{31}x_{t1}+\beta_{32}x_{t2}+\beta_{33}x_{t3}+\beta_{34}x_{t4}+\varepsilon_{t3}$$
$$y_{t4}=\gamma_{41}y_{t1}+\gamma_{42}y_{t2}+\gamma_{43}y_{t3}+\beta_{41}x_{t1}+\beta_{42}x_{t2}+\beta_{43}x_{t3}+\beta_{44}x_{t4}+\varepsilon_{t4}$$

فرم حل شده مدل فوق بصورت زیر خواهد بود:

$$Y = X\Pi + V$$

به طوریکه:

$$V = E(I-\Gamma)^{-1} \qquad \Pi = B(I-\Pi)^{-1}$$

Π ماتریس ۴,۴ است. بنابراین داریم:

$$y_{t1}=\pi_{11}x_{t1}+\pi_{12}x_{t2}+\pi_{13}x_{t3}+\pi_{14}x_{t4}+v_{t1}$$
$$y_{t2}=\pi_{21}x_{t1}+\pi_{22}x_{t2}+\pi_{23}x_{t3}+\pi_{24}x_{t4}+v_{t2}$$
$$y_{t3}=\pi_{31}x_{t1}+\pi_{32}x_{t2}+\pi_{33}x_{t3}+\pi_{34}x_{t4}+v_{t3}$$
$$y_{t4}=\pi_{41}x_{t1}+\pi_{42}x_{t2}+\pi_{43}x_{t3}+\pi_{44}x_{t4}+v_{t4}$$

ملاحظه میشود که در فرم خلاصه شده بالا همه متغیرهای درونزا بر حسب کلیه متغییرهای از پیش تعیین شده محاسبه شده‌اند و ضرایب فرم خلاصه شده π ها بر حسب پارامترهای فرم ساختاری هستند.

فرم خلاصه شده بالا با استفاده از روش حداقل مربعات معمولی قابل تخمین است و تخمین زننده‌های سازگاری بدست میآید. اما قادر نخواهیم بود بادر دست داشتن تخمین پارامترهای فرم خلاصه شده فوق پارامترهای فرم ساختاری موجود را شناسایی کنیم. لذا هیچیک از معادلات مدل فوق قابل شناسایی نیستند زیرا بدون اعمال قید در فرم ساختاری پارامترها قابل شناسایی نخواهند بود.

حال اجازه دهید برای شناسایی معادله اول متغیر y_{t4} ومتغیرهای x_{t3} , x_{t4} را از معادله اول حذف کنیم. پس قیدهای زیر را اعمال میکنیم.

$\gamma_{14}=0$, $\beta_{13} = \beta_{14}=0$

در این صورت فرم خلاصه شده را میتوانیم بصورت زیر بنویسیم:

$$\begin{bmatrix} y_{.1} & Y_1 & Y_1^* \end{bmatrix} = \begin{bmatrix} X_1 & X_1^* \end{bmatrix} \begin{bmatrix} \pi_{11} & \pi_{21} & \pi_{31} & \pi_{41} \\ \pi_{12} & \pi_{22} & \pi_{32} & \pi_{42} \\ \pi_{13} & \pi_{23} & \pi_{33} & \pi_{43} \\ \pi_{14} & \pi_{24} & \pi_{34} & \pi_{44} \end{bmatrix} + \begin{bmatrix} v_1 & V_1 & V_1^* \end{bmatrix}$$

$$Y_1 = \begin{bmatrix} y_{t2} & y_{t3} \end{bmatrix} \qquad Y_1^* = \begin{bmatrix} y_{t4} \end{bmatrix}$$

$$X_1 = \begin{bmatrix} x_{t1} & x_{t2} \end{bmatrix} \qquad X_2^* = \begin{bmatrix} x_{t3} & x_{t4} \end{bmatrix}$$

$$Y_1 = X_1 \Pi_{12} + X_1^* \Pi_{22} + V_1$$

در رابطه اخیر:

$$\Pi_{12} = \begin{bmatrix} \pi_{21} & \pi_{31} \\ \pi_{22} & \pi_{32} \end{bmatrix} \qquad \Pi_{22}^* = \begin{bmatrix} \pi_{23} & \pi_{33} \\ \pi_{24} & \pi_{34} \end{bmatrix}$$

نمادهای ستاره‌دار نشانگر این معناست که این متغیرها (دسته‌ای از متغیرها) در معادله اول ظاهر نشده اند.

از فرم خلاصه شده , میدانیم که:

$$\Pi(I - \Gamma) = \mathrm{B}$$

وبرای معادله اول خواهیم داشت:

$$\begin{bmatrix} \pi_{11} & \pi_{21} & \pi_{31} & \pi_{41} \\ \pi_{12} & \pi_{22} & \pi_{32} & \pi_{42} \\ \pi_{13} & \pi_{23} & \pi_{33} & \pi_{43} \\ \pi_{14} & \pi_{24} & \pi_{34} & \pi_{44} \end{bmatrix} \begin{bmatrix} 1 \\ -\gamma_{12} \\ -\gamma_{13} \\ 0 \end{bmatrix} = \begin{bmatrix} \beta_{11} \\ \beta_{12} \\ 0 \\ 0 \end{bmatrix}$$

ویا میتوانیم دستگاه زیر را از روابط بالابدست آوریم.

$$\pi_{11} - \gamma_{12}\pi_{21} - \gamma_{13}\pi_{31} + 0.\pi_{41} = \beta_{11}$$
$$\pi_{12} - \gamma_{12}\pi_{22} - \gamma_{13}\pi_{32} + 0.\pi_{42} = \beta_{12}$$
$$\pi_{13} - \gamma_{12}\pi_{23} - \gamma_{13}\pi_{33} + 0.\pi_{43} = 0$$
$$\pi_{14} - \gamma_{12}\pi_{24} - \gamma_{13}\pi_{34} + 0.\pi_{42} = 0$$

در اینجا برای آنکه بتوانیم پارامترهای فرم ساختاری (γها) را از دستگاه بالا بر حسب پارامترهای فرم خلاصه شده (π ها) بدست آوریم بصورت زیر عمل میکنیم:

$$\gamma_{12}\pi_{21}+\gamma_{13}\pi_{31}+\beta_{11}=\pi_{11}$$
$$\gamma_{12}\pi_{22}+\gamma_{13}\pi_{32}+\beta_{12}=\pi_{12}$$
$$\gamma_{12}\pi_{23}+\gamma_{13}\pi_{33}=\pi_{13}$$
$$\gamma_{12}\pi_{24}+\gamma_{13}\pi_{34}=\pi_{14}$$

لذا از معادله سوم و چهارم ابتدا γ ها را بدست میآوریم و داریم:

$$\begin{bmatrix}\pi_{23} & \pi_{33}\\ \pi_{24} & \pi_{34}\end{bmatrix}\begin{bmatrix}\gamma_{12}\\ \gamma_{13}\end{bmatrix}=\begin{bmatrix}\pi_{13}\\ \pi_{14}\end{bmatrix}\Rightarrow\begin{bmatrix}\gamma_{12}\\ \gamma_{13}\end{bmatrix}=\begin{bmatrix}\pi_{23} & \pi_{33}\\ \pi_{24} & \pi_{34}\end{bmatrix}^{-1}\begin{bmatrix}\pi_{13}\\ \pi_{14}\end{bmatrix}$$

بنابراین برای بدست آوردن معکوس ماتریس ۲,۲ نیاز داریم که مرتبه ماتریس زیربرابر ۲ باشد:

$$\Pi_{22}=\begin{bmatrix}\pi_{23} & \pi_{33}\\ \pi_{24} & \pi_{34}\end{bmatrix}$$

یعنی: $\mathrm{Rank}(\Pi_{22})$ باید۲ باشد.

تعداد متغیرهای درونزای ظاهر شده در معادله اول $G_1=۲$ میباشد. از طرفی تعداد متغیرهای برونزای ظاهر نشده$K_1^*=K-K_1$ در معادله اول برابر ۲ است.

$$K_1^*=G_1$$

پس معادله یکم دقیقا قابل شناسایی است.

شرط درجه‌ای: شرطی است لازم و نه کافی برای معادله i ام و زمانی برقرار است که:

$K_i^*\geq Gi$ یعنی تعداد متغیرهای برونزای کنار گذاشته شده از معادله i ام حداقل برابر تعداد متغیرهای درونزای وارد شده در معادله i ام باشد. این شرط صرفا شرطی شمردنی است و اگر برقرار نباشد نیازی به بررسی شرط رتبه ای نیست.

شرط رتبه‌ای: شرطی است لازم وکافی برای شناسایی معادله i ام وزمانی برقرار است که:

$K_i^*=G_i$ یعنی تعداد متغیرهای برونزای کنار گذاشته شده از معادله i ام برابر تعداد متغیرهای درونزای وارد شده در معادله i ام باشد.

ملاحظه میفرمایید که تنها با خارج کردن یک متغیردرونزا و دو متغیر برونزا از معادله اول قادر به قابل شناسایی نمودن پارامترهای این معادله گردیدیم.

تخمین به روش OLS برای فرم خلاصه شده و فرم ساختاری:

الف- OLS برای فرم ساختاری:

$$\text{Lemma1.}\quad P\lim \frac{U'U}{T} = \Sigma$$

$$\text{i.e}\quad P\lim \frac{1}{T}\Sigma u_{ti}u_{tj} = \sigma_{ij}$$

$$\text{Lemma2}\quad .P\lim \frac{V'V}{T} = \Omega$$

یـــادآوری:

$$P\lim \frac{Z'U}{T} = P\lim \frac{Z'V}{T} = 0$$

$$\text{Lemma3.}\quad P\lim \frac{Y'U}{T} = (I - B')^{-1}\Sigma \neq 0$$

اثبات:

$$P\lim \frac{Y'U}{T} = P\lim\left[\pi' \frac{Z'U}{T} + (I-B')^{-1}\frac{U'U}{T}\right] = (I-B')^{-1}\Sigma$$

چون جمله داخل کروشه برابر صفر است.

قضیه: اگر تخمین OLS پارامترهای معادله i ام از فرم ساختاری را $\delta_{i,OLS}$ بنامیم.

$$\delta_{i,OLS} = (X_i'X_i)^{-1}X_i'Y_{.i} \qquad X_i = [Y,Z]L_i$$

آیا OLS سازگار است؟

باید از آن $P\lim$ گرفته شود.

$$P\lim \delta_{i,OLS} = P\lim \delta_i + P\lim\left(\frac{X_i'X_i}{T}\right)^{-1}\frac{X_i'u_{.i}}{T}$$

$$= \delta_i + P\lim\left(\frac{X_i'X_i}{T}\right)^{-1} P\lim\left(\frac{X_i'u_{.i}}{T}\right)$$

داریم :

$$X_i'X_i = L_i'\begin{bmatrix} Y' \\ Z' \end{bmatrix}[Y,Z]L_i$$

$$= L_i'\begin{bmatrix} Y'Y & Y'Z \\ Z'Y & Z'Z \end{bmatrix}L_i$$

$$X_i'u_{.i} = L_i'\begin{bmatrix} Y'u_{.i} \\ Z'u_{.i} \end{bmatrix}$$

$$P\lim \frac{X_i'X_i'}{T} = L_i'\begin{bmatrix} P\lim \frac{YY}{T} & P\lim \frac{Y'Z}{T} \\ P\lim \frac{Z'Y}{T} & P\lim \frac{Z'Z}{T} \end{bmatrix}L_i$$

$$1) P\lim \frac{Z'Z}{T} = Q$$

$$2)P\lim\frac{Y'Y}{T}=P\lim\frac{1}{T}\left[Z\prod+V\right]'\left[Z\prod+V\right]$$
$$=P\lim\frac{1}{T}\left[\prod' Z'Z\prod+V'Z\prod+\Pi' Z'V+V'V\right]$$
$$=P\lim\prod'\frac{Z'Z}{T}\prod+\prod P\lim\frac{V'Z}{T}+\prod' P\lim\frac{Z'V}{T}+P\lim\frac{V'V}{T}$$
$$=\prod' Q\prod+\Omega$$

$$3)P\lim\frac{Y'Z}{T}=P\lim\frac{1}{T}\left[Z\prod+V\right]' Z$$
$$=\prod' P\lim\frac{Z'Z}{T}+P\lim\frac{V'Z}{T}$$
$$=\prod' Q$$

$$P\lim\frac{X'_{.i}u_{.i}}{T}=L'_i\begin{bmatrix}P\lim\dfrac{Y'u_{.i}}{T}\\ P\lim\dfrac{Z'u_{.i}}{T}\end{bmatrix}$$

چون:

$$P\lim\frac{Y'U}{T}=\left(I-B\right)^{-1}\sum\Rightarrow\left(I-B'\right)^{-1}\sigma^{.i}$$

چون:

$$P\lim\frac{Z'U}{T}=0\Rightarrow P\lim\frac{Z'u_{.i}}{T}=0$$

لذا:

$$P\lim\frac{X'_i u_{.i}}{T}=L'_i\begin{bmatrix}\left(I-B'\right)^{-1}\sigma^i\\ 0\end{bmatrix}$$

در نتیجه خواهیم داشت:

$$\Rightarrow P\lim\delta_{i,OLS}=\delta_i+\left\{L'_i\begin{bmatrix}\prod' Q\prod+\Omega & \prod' Q\\ Q'\Pi & Q\end{bmatrix}L_i\right\}^{-1}\left\{L'_i\begin{bmatrix}\left(I-B'\right)^{-1}\sigma^{.i}\\ 0\end{bmatrix}\right\}\neq\delta_i$$

پس تخمین‌زننده های OLS بطورکلی برای فرم ساختاری معادلات همزمان ناسازگار ند و این به دلیل مشکل وجود رابطه بین Y_i و $U_{.i}$ می‌باشد.

ب-کاربرد روش OLS در فرم خلاصه شده:

$$y=Z\pi+v\quad;\quad y=Vec(Y)$$

بطوریکه:

$$v=Vec(V)$$
$$y=\left(I_G\otimes Z\right)vec\left(\prod\right)+v$$

یا:

$$\begin{bmatrix} y_{.1} \\ \vdots \\ y_{.G} \end{bmatrix} = \begin{bmatrix} Z & & 0 \\ & \ddots & \\ 0 & & Z \end{bmatrix} \begin{bmatrix} \pi_{.1} \\ \vdots \\ \pi_{.G} \end{bmatrix} + \begin{bmatrix} v_{.1} \\ \vdots \\ v_{.G} \end{bmatrix}$$

توجه داریم که فرم خلاصه شده شبیه مدل SUR است بطوریکه:

$$E(v) = 0, \quad E(vv') = \Omega \otimes I_T$$

تحت شرایط نرمال بودن جملات اخلال تخمین GLS (یعنی SUR) و ML یکی خواهد بود. از طرفی چون دارای یک مجموعه از متغیرهای مستقل هستیم، لذا SUR تبدیل به OLS خواهد شد.

لذا تخمین OLS پارامترهای فرم خلاصه شده π و Ω عبارتسند از:

$$\hat{\Pi} = (ZZ')^{-1} Z'Y$$

$$\Omega = \frac{1}{T} \hat{V}' \hat{V}, \quad \hat{V} = Y - Z\hat{\Pi}$$

$$P\lim \hat{\Pi} = \Pi + P\lim \left(\frac{Z'Z}{T} \right)^{-1} P\lim \frac{Z'V}{T} = Q.0 = 0$$

$$P\lim \hat{\Pi} = \Pi$$

ملاحظه می‌شود که $\hat{\Pi}$ تخمین‌زننده‌ای سازگار است:

$$\sqrt{T}\ \mathrm{vec}\left(\hat{\Pi} - \Pi \right) \xrightarrow{1.d} N(0, \quad \Omega \otimes \bar{Q})$$

و اگر U نرمال باشد، $\hat{\Pi}$ و $\hat{\Omega}$ تخمین‌زننده‌های ML از Π و Ω هستند.

۱۳-۷ حداقل مربعات غیرمستقیم:

جریان روش تخمین حداقل مربعات غیرمستقیم بدین ترتیب است که ما ابتدا پارامترهای فرم حل شده را تخمین می‌زنیم سپس بر اساس آن تخمین‌زننده‌های ابتدایی پارامترهای فرم ساختاری یعنی B و C را تخمین خواهیم زد.

همانطور که گفته شد ما رابطه زیر را بین پارامترهای مدل حل شده و ساختاری داریم

$$\Pi = C(I - B)^{-1}$$

$$\Pi(I - B) = C$$

$$\Pi - \Pi B = C$$

$$\Pi B + C = \Pi$$

$$[\Pi, \quad I_K]\begin{bmatrix} B \\ C \end{bmatrix} = \Pi$$

برای معادله i ام خواهیم داشت:

$$[\Pi, \quad I_K]\begin{bmatrix} b_{.i} \\ c_{.i} \end{bmatrix} = \Pi_{.i}$$

$$(a) \qquad [\pi_{.i}, L_{i2}]\begin{bmatrix} \beta_i \\ \gamma_i \end{bmatrix} = \Pi_{.i} \qquad \Pi_i = \Pi L_{i1}$$

نیاز داریم که مرتبه ماتریس (Π_i, L_{i2}) برابر با $G_i + K_i$ باشد.

حال بعنوان مثال معادله اول را در نظر بگیریم:

$$y_{.1} = Y_1\beta_1 + Z_1\gamma_1 + u_{.1}$$

$$Y = [y_{.1}, Y_1, Y_1^*], \quad Z = [Z_1, Z_1^*], \quad \Pi = \begin{bmatrix} \Pi_{11} & \Pi_{12} & \Pi_{13} \\ \Pi_{21} & \Pi_{22} & \Pi_{23} \end{bmatrix}$$

آنگاه:

$$L_{12} = \begin{bmatrix} I_{K_1} \\ O_{(K-K_1)\times k_1} \end{bmatrix} \quad \text{و} \quad \Pi_1 = \Pi L_{11} = \begin{bmatrix} \Pi_{12} \\ \Pi_{22} \end{bmatrix}$$

لذا می‌توانیم رابطه a (که قبلاً هم دیدیم) را بصورت زیر بنویسیم:

$$\Pi_{12}\beta_1 + \gamma_1 = \Pi_{11}$$

$$\Pi_{22}\beta_1 \quad = \Pi_{12}$$

برای مسئله تشخیص ما نیاز داشتیم که $Rank(\Pi_{22}) = k_1$ باشد.

یا:

$$Rank\begin{bmatrix} \Pi_{12} & I_{K_1} \\ \Pi_{22} & 0 \end{bmatrix} = G_1 + K_1$$

در اینجا متوجه می‌شویم که نیاز ما برای تشخیص در رابطه a این است که: $Rank(\Pi_i, L_{12}) = G_i + K_i$ باشد.

اگر $K = G_i + K_i$ باشد معادله i ام دقیقاً قابل تشخیص و اگر $K > G_i + K_i$ باشد فوق تشخیص خواهد بود.

تعریف ۱ – فرض کنید $\hat{\Pi}$ ، تخمین OLS پارامترهای فرم حل شده باشد.

i) در حالتی که معادله i ام دقیقاً قابل تشخیص باشد، ILS را بصورت بردار جوابهای منحصر بفرد زیر تعریف می‌کنیم.

$$\left[\hat{\Pi}_i, L_{i2}\right]\delta_{i,OLS} = \hat{\Pi}_{.i}$$

ii) در مواردی که معادله یکم فوق تشخیص است $K > G_i + K_i$، ماتریس S با ابعاد $(G_i + K_i) \times K$ را بعنوان یک ماتریس انتخابگر در نظر می‌گیریم بطوریکه با پیش ضرب کردن آن در ماتریس‌های فوق، $K - (G_i + K_i)$ سطر را حذف می‌کنیم.

سپس تخمین‌زننده ILS (متناظر با S) را بصورت زیر تعریف می‌کنیم.

$$S\left[\hat{\Pi}_i, L_{i2}\right]\delta_{i,ILS} = S\,\Pi_{.i}$$

نکته ۱ – فرض کنید می‌خواهیم $K - G_i - K_i$ سطر را حذف کنیم، آنگاه:

$$S = \left[I_{G_i+K_i}, O_{(G_i+K_i)\times(K-G_i-K_i)}\right]$$

توجه داریم که برای پارامترهای فرم حل شده واقعی، جواب این دستگاه:

$$\left[\Pi_i, L_{i2}\right]\delta_i = \Pi_1$$

بدون توجه به سطوری که ما حذف می‌کنیم، یکی خواهد شد.

ولی، وقتی π_i را با تخمین‌زننده OLS، $\hat{\pi}_i$ جایگزین می‌کنیم، بطورکلی جوابهای مختلفی خواهیم داشت و آن جوابها بستگی دارند به اینکه کدام سطرها را حذف می‌کنیم.

نکته۲ – بطورکلی در این حالت که معادله فوق تشخیص است ILS تخمین‌زننده کارآیی نخواهد بود. دلیل آن این است که ما بخشی از اطلاعات را نادیده گرفته‌ایم.

قضیه : تخمین‌زننده IV (ابزاری) زیر را در نظر بگیریم:

$$\delta_{i,IV} = \left(\hat{X}_i' X_i\right)^{-1} \hat{X}_i' y_{.i}$$

$$\hat{X}_i = \left[\hat{Y}_i, Z_i\right]$$

$$\hat{Y}_i = Z\,\hat{\Pi}_i$$

$$\hat{\Pi}_i = \hat{\Pi} L_{i1}$$

$$\hat{\Pi} = (Z'Z)^{-1} Z'Y$$

یعنی متغیر ابزاری ما بر اساس تخمین OLS فرم خلاصه شده بدست آمده است.

حالا فرض کنید معادله iام دقیقاً قابل تشخیص باشد، آنگاه IV یعنی $\delta_{i,IV}$ و ILS به لحاظ عددی برابرند. بعداً این IV را $2SLS$ خواهیم نامید.

اثبات: چون معادله iام دقیقاً قابل تشخیص است لذا $S = I$ خواهد بود و تخمین‌زننده ILS بصورت زیر است:

$$\left[\hat{\Pi}_i, L_{i2}\right]\delta_{i,ILS} = \hat{\Pi}_{.i}$$

$$\begin{bmatrix} Y' \\ Z' \end{bmatrix}\left[\hat{\Pi}_i, L_{i2}\right]\hat{\delta}_{i,ILS} = \begin{bmatrix} Y_i' \\ Z_i' \end{bmatrix}\hat{\Pi}_{.i}$$

حال[۷۱] :

$$\hat{\prod}_i = \hat{\prod} L_{i1} = (Z'Z)^{-1} Y_i$$
$$L_{i2} = (Z'Z)^{-1} Z'ZL_{i2} = (Z'Z)^{-1} Z'Z_i$$
$$\hat{\prod}_{.i} = (Z'Z)^{-1} Z'y_{.i}$$

لذا:

$$\begin{bmatrix} Y_i' \\ Z_i' \end{bmatrix} \left\{ (Z'Z)^{-1} Z'Y_i, \quad Z(Z'Z)^{-1} Z'Z_i \right\} = \begin{bmatrix} Y_i \\ Z_i' \end{bmatrix} Z(Z'Z)^{-1} Z'[Y_i, Z_i]\delta_{i,ILS}$$

یا:

$$= \begin{bmatrix} Y_i' \\ Z_i' \end{bmatrix} Z(Z'Z)^{-1} Z'y_{.i}$$

داریم:

$$\begin{bmatrix} \hat{Y}_i' \\ Z_i' \end{bmatrix} [Y_i, Z_i]\delta_i = \begin{bmatrix} \hat{Y}_i' \\ Z_i' \end{bmatrix} y_{.i}$$
$$\left(\hat{X}_i' X_i \right) \left(\hat{X}_i' y_i \right)^{-1} \hat{X}_i' y_{.i} = \hat{X}_i' y_{.i}$$

اثبات شد:

$$\hat{X}_i' y_{.i} = \hat{X}_i' y_{.i}$$

۱۳-۸ تخمین زننده با اطلاعات محدود[۷۲]:

سه تعبیر برای 2SLS :

الف - $2SLS$ بعنوان متغیر ابزاری:

کلیه متغیرهای درونزایی که در معادله i ام ظاهر شده‌انـد را روی همـه متغیرهـای از پـیش تعیـین شده رگرس می‌کنیم و این عبارتست از:

$$\hat{\prod}_{i,OLS} = (Z'Z)^{-1} Z'Y_i$$
$$\hat{Y}_i = Z \hat{\prod}_{i,OLS} = Z(Z'Z)^{-1} Z'Y_i$$

متغیر ابزاری:

$$\hat{X}_i = \left[\hat{Y}_i, Z_i \right]$$

چون:

$$\hat{X}_i' u_{.i} = 0$$
$$\hat{X}_i' y_{.i} = \hat{X}_i' X_i \delta_i + \hat{X}_i U_{.i}$$

(۱)

71 . $Z_i'Z(Z'Z)^{-1} Z', L_{i2}'Z'Z(Z'Z)^{-1} Z', L_{i2}'Z' = Z_i'$

72 . Limited Information Estimator

$$\hat{\delta}_{i_{2SLS}} = \left(\hat{X}_i' X_i \right)^{-1} \hat{X}_i' y_{.i}$$

ب- استفاده از OLS در دو مرحله:

قدم اول- فرم خلاصه شده را با OLS تخمین می‌زنیم.

$$\hat{\prod}_{i,OLS} = (Z'Z)^{-1} Z'Y_i$$

$$\hat{Y}_i = Z \hat{\prod}_{i,OLS} = Z(Z'Z)^{-1} Z'Y_i$$

قدم دوم - Y_i های معادله i ام را با $\hat{Y}_i$ جایگزین می‌کنیم.

$$y_{.i} = \hat{Y}_i \beta_i + Z_i \gamma_i + u_{.i} = \hat{X}_i \delta_i + u_{.i}$$

(۲)

$$\hat{\delta}_{i,2SLS} = \left(\hat{X}_i' \hat{X}_i \right)^{-1} \hat{X}_i' u_{.i}$$

نشان می‌دهیم که : ۲=۱

اثبات:

$$(۱) \quad \hat{\delta}_{i,2SLS} = \begin{bmatrix} \hat{Y}_i'Y & \hat{Y}_i'Z_i \\ Z_i'Y_i & Z_i'Z_i \end{bmatrix}^{-1} \begin{bmatrix} \hat{Y}_i' & y_{.i} \\ Z_i' & y_{.i} \end{bmatrix}$$

$$(۲) \quad \hat{\delta}_{i,2SLS} = \begin{bmatrix} \hat{Y}_i' \hat{Y}_i & \hat{Y}_i' Z_i \\ Z_i' \hat{Y}_i & Z_i' Z_i \end{bmatrix}^{-1} \begin{bmatrix} \hat{Y}_i' y_{.i} \\ Z'_{i\,y.i} \end{bmatrix}$$

اگر بتوان نشان داد که $\hat{Y}_i'\hat{Y}_i = \hat{Y}_i'Y_i$ و $Z_i'\hat{Y}_i = Z_i'Y_i$ قضیه ثابت است:

$$\hat{Y}_i'\hat{Y}_i'Z(Z'Z)^{-1} Z'Y_i$$

داریم:

$$= Y_i'Z(Z'Z)^{-1} Z'Y_i = \hat{Y}' Y_i$$

لذا خواهیم داشت:

$$Z_i'\hat{Y}_i = L_{i2}'Z'Z(Z'Z)^{-1} Z'Y_i = L_{i2}'Z'Y_i = Z_i'Y_i$$

پس رابطه (۱) و (۲) برابر است.

ج- 2SLS بعنوان حداقل مربعات تعمیم یافته[۷۳] :

$$y_{.i} = X_i \delta_i + u_{.i}$$

73 . Generalized Least Square

اگر در این مدل Z' را بعنوان ماتریس تبدیل مدل بکار ببریم بطوریکه می‌دانیم:

$$EZ'u_{.i}u'_{.j}Z = \sigma_{ij}Z'Z$$

آنگاه:

$$\hat{\delta}_{i,2SLS} = \left[X_i'Z(Z'Z)^{-1}Z'X_i\right]^{-1}X_i'Z(Z'Z)^{-1}Z'y_{.i}$$

چون در این تخمین‌زننده‌ها $\hat{\Pi}_{i,OLS}$ بکار برده شده یک تخمین‌زننده سازگار برای $\hat{\Pi}_i$ است، قضایای مربوط به حدی مرکزی را راجع به این تخمین‌زننده‌ها $(LIVE)$ [۷۴] می‌توانیم بکار بریم:

بطوریکه:

$$\hat{\delta}_{i,IILS} \sim N\left(\delta_i, \tilde{\Phi}_i\right)$$

$$\tilde{\Phi}_i = \tilde{\sigma}_{ii}\left(\tilde{X}_i'\tilde{X}_i\right)^{-1}$$

$$\hat{\sigma}_{ii} = \frac{1}{T}\left[y_{.i} - X_i\delta_{i,2SLS}\right]'\left[y_{.i} - X_i\delta_{i,2SLS}\right]$$

قضیه: تخمین‌زننده K-class یک تخمین‌زننده IV می‌باشد؛ $Y_1 - K\hat{V}_1$ یک متغیر ابزاری برای Y_1 و Z_1 متغیر ابزاری خودش است.

اثبات: می‌دانیم $\left(Y_1 - K\hat{V}_1\right)'Z_1 = Y_1'Z_1$

چون : $\hat{V}_1'Z = 0$

همچنین $\left(Y_1 - K\hat{V}_1\right)'Y_1 = Y_1'Y_1 - K\hat{V}_1'Y_1$

چون : $Y_1 = \hat{Y}_1 + \hat{V}_1$ ؛ $= Y_1'Y_1 - K\hat{V}_1'\hat{V}_1 - K\hat{V}_1\hat{Y}_1$

چون : $\hat{V}_1'\hat{Y}_1 = 0$ ؛ $= Y_1'Y_1 - K\hat{V}_1'\hat{V}_1$

آنگاه تخمین زننده IV برای γ و β عبارتست از :

$$\begin{bmatrix}\hat{\beta}_i\\ \hat{\gamma}_i\end{bmatrix} = \begin{bmatrix}\left(Y_1 - K\hat{V}_1\right)' & \left(Y_1 - K\hat{V}_1\right)'Z_1\\ Z_1'Y_1 & Z_1'Z_1\end{bmatrix}^{-1}\begin{bmatrix}\left(Y_1 - K\hat{V}_1\right)'u_{.i}\\ Z_1'u_{.i}\end{bmatrix} = \begin{bmatrix}\hat{\beta}_i\\ \hat{\gamma}_i\end{bmatrix}_{k-class}$$

قضیه: اگر $P\lim k = 1$ باشد تخمین‌زننده‌های K-class سازگارند و بطور مجانبی دارای ماتریس واریانس – کوواریانس 2SLS است.

اگر $P\lim\sqrt{T}(K-1) = 0$، البته اگر معادله قابل تشخیص است[۷۵].

74 . Limited Instrumental Variables Estimator

75 ، ۱۶۷ اثبات صفحه .Peteer. Schmidet

تعریف: تخمین زننده γ و β در معادله i ام.

$$Y_{.i} = Y_1\beta_i + Z_1\gamma_i + u_{.i}$$

عبارتست از :

$$\begin{bmatrix} \hat{\beta}_i \\ \hat{\gamma}_i \end{bmatrix} = \begin{bmatrix} Y_1'Y_1 - K\hat{V}'\hat{V}_1 & Y_1'Z_1 \\ Z_1'Y_1 & Z_1'Z_1 \end{bmatrix}^{-1} \begin{bmatrix} \left(Y_1 - K\hat{V}_1\right)' y_{.i} \\ Z_1'y_{.i} \end{bmatrix}$$

بطوریکه:

$$\hat{V}_1 = Y_1\hat{Y}_1 = Y_1 - I\left(Z'Z\right)^{-1}Z'Y_1$$

قضیه (۱) :

تخمین‌زننده OLS یک تخمین‌زننده k-class است بطوریکه k برابر صفر است.

قضیه (۲) :

تخمین‌زننده 2SLS یک تخمین‌زننده k-class است بطوریکه k=1 است.

اثبات:

$$\hat{Y}_1 = Y_1 - \hat{V}_1$$

در نتیجه :

$$\left(\hat{Y}_1 - (1)\hat{V}_1\right)' Y = \hat{Y}_1'Y$$

هم‌چنین:

$$Y_1'Y_1 - (1)\hat{V}_1'\hat{V}_1 = Y_1'Y_1 - \left(Y_1 - \hat{Y}_1\right)'\left(Y_1 - \hat{Y}_1\right)$$

$$= Y_1'Y_1 - Y_1'Y_1 + Z\hat{Y}_1'Y_1 - \hat{Y}_1'\hat{Y}_1$$

$$= \hat{Y}'_1 Y_1$$

چون $\hat{Y}_1'Y_1 = Y_1'\hat{Y}_1 = \hat{Y}_1'\hat{Y}_1$ بنابراین:

$$\begin{bmatrix} \hat{\beta}_i \\ \hat{\gamma}_i \end{bmatrix} = \begin{bmatrix} \hat{Y}_1'Y_1 & Y_1'Z_1 \\ Z_1'Y_1 & Z_1'Z_1 \end{bmatrix}^{-1} \begin{bmatrix} \hat{Y}_1'y_{.i} \\ Z_1'u_{.i} \end{bmatrix} = \begin{bmatrix} \hat{\beta}_i \\ \hat{\gamma}_i \end{bmatrix}_{2SLS}$$

۱۳-۹ تخمین‌زننده با اطلاعات کامل[۷۶] :

میدانیم که:

$$\tilde{Y} = Z\tilde{\prod}_{ols}$$

---

76 . Full Information Instrumental Estimator

$$\tilde{\Pi}_{OLS} = (Z'Z)^{-1} Z'Y$$
$$\tilde{\Sigma}_{2SLS} = \left(\tilde{\delta}_{ij,2SLS}\right)$$
$$\tilde{\delta}_{i,j,2SLS} = \frac{1}{T}\left[y_{.i} - X_i \delta_{i,2SLS}\right]'\left[y_j - X_j \delta_{i,2SLS}\right]$$

براساس روابط فوق تخمین‌زننده 3SLS را بصورت زیر تعریف می‌کنیم.

الف -

$$\hat{\delta}_{3SLS} = \left[\tilde{X}'_{OLS}\left(\tilde{\Sigma}^{-1}_{2SLS} \otimes I_T\right)\tilde{X}_{OLS}\right]^{-1} \tilde{X}'_{OLS}\left(\tilde{\Sigma}^{-1}_{2SLS} \otimes I_T\right) y$$

$$= \begin{bmatrix} \tilde{\sigma}^{11} \tilde{X}'_1 X_1 \ldots \tilde{\sigma}^{1G} \tilde{X}'_1 X_G \\ \vdots \\ \tilde{\sigma}^{G1} \tilde{X}'_G X_1 \ldots \tilde{\sigma}^{GG} \tilde{X}_G{}' X_G \end{bmatrix}^{-1} \begin{bmatrix} \sum_{i=1}^{G} \tilde{\sigma}^{1i} \tilde{X}'_1 y_{.i} \\ \vdots \\ \sum_{i=1}^{G} \tilde{\sigma}^{Gi} X'_G y_{.i} \end{bmatrix}$$

$$(1)\ \tilde{Y}'_i \tilde{Y}_j = Y'_i Z(Z'Z)^{-1} Z'Z(Z'Z)^{-1} Z'Y_j = Z'_j Y_j = Y'_j Z(Z'Z)^{-1} Z'Y_j$$
$$\tilde{Y}'_i Z_j = Y'_i Z(Z'Z_j) = Y'_i Z_j \qquad ;(Z_j = ZL_{i2})$$

$$(2)\ Z'_i Z_j = Z'_j Z(Z'Z)^{-1} Z'Z_j = Z'Z_j = Z'_i Z(Z'Z)^{-1} Z'ZL_{i2} = Z'_i Z_j$$

از (۱) بدست می‌آید که :

$$(3)\ \tilde{X}'_i X_j = \begin{bmatrix} \tilde{Y}'_i Y_j & \tilde{Y}'_i Z_j \\ Z_i Y_j & Z'_i Z_j \end{bmatrix} = \begin{bmatrix} \tilde{Y}'_i \tilde{Y}_j & \tilde{Y}'_i Z_j \\ Z'_i \tilde{Y}_j & Z'_i Z_j \end{bmatrix} = \tilde{X}' \tilde{X}_j$$

رابطه اخیر به ما اجازه می‌دهد که متغیر دومی برای 3SLS بدست آوریم.

ب- تخمین زننده‌ای که دو بار OLS بکار ببریم و سپس حداقل مربعات وزنی[۷۷] :

قدم اول : π را از فرم خلاصه شده با OLS تخمین می‌زنیم و $\tilde{Y}$ را بدست می‌آوریم.

$$\tilde{Y} = Z(Z'Z)^{-1} Z'Y \qquad \text{یا} \qquad \tilde{Y}_i = \tilde{Y} L_{i1} \quad i = 1, \ldots G$$

قدم دوم : تخمین OLS معادله $y_{.i} = \tilde{Y}_i \beta_i + Z_i \gamma_i + u_{.i} \quad i = 1, \ldots, G$

که 2SLS به ما خواهد داد: یعنی $\delta_i \quad i = 1, \ldots, G$

77 . Weighted Least Square

از این 2SLS استفاده می‌کنیم و Σ را تخمین می‌زنیم:

$$\tilde{\sum}=\left[\tilde{\sigma}_{ii,2SLS}\right]$$

$$\tilde{\sigma}_{ij,2SLS}=\frac{1}{T}\left[y_{.i}-X_i\delta_{i,2SLS}\right]'\left[y_{.j}-X_j\delta_{j,2SLS}\right]$$

قدم سوم: حداقل مربعات وزنی را در مدل زیر بکار می‌بریم:

$$y=\tilde{X}_{OLS}\,\delta+\varepsilon$$

$$\varepsilon=u+\left(X-\tilde{X}_{OLS}\right)\delta$$

ماتریس وزنی عبارتست از :

$$\Sigma_{2SLS}^{-1}\otimes I_T$$

اما :

$$Euu'=\Sigma\otimes I_T$$

$$E(\varepsilon\varepsilon')\neq\Sigma\otimes I_T$$

آنگاه داریم:

$$\delta_{3SLS}=\left[\tilde{X}'_{OLS}\left(\tilde{\Sigma}_{2SLS}^{-1}\otimes I_T\right)\tilde{X}_{OLS}\right]^{-1}\tilde{X}'_{OLS}\left(\tilde{\Sigma}_{2SLS}^{-1}\otimes I_T\right)y \qquad \text{(ب)}$$

$$\delta_{3SLS}=\begin{bmatrix}\tilde{\sigma}^{11}\tilde{X}_1'\tilde{X}_1 & \dots & \tilde{\sigma}^{1G}\tilde{X}_1'\tilde{X}_G\\ \vdots & & \dots\\ \tilde{\sigma}^{G1}\tilde{X}_G'\tilde{X}_1 & \dots & \tilde{\sigma}^{GG}\tilde{X}_G'X_G\end{bmatrix}^{-1}\begin{bmatrix}\sum_{i=1}^{G}\tilde{\sigma}^{1i}\tilde{X}_1'y_{.i}\\ \vdots\\ \sum_{i=1}^{G}\tilde{\sigma}^{Gi}\tilde{X}'_G\,y_{.i}\end{bmatrix}$$

بطوریکه $y=Vec(Y)$ و $u=Vec(U)$ می‌باشند و می‌توان نشان داد که رابطه "الف" و "ب" با یکدیگر برابرند. چون : $\tilde{X}'\tilde{X}_j=\tilde{X}_i'X_j$

فصل چهاردهم

تحلیل مدل سریهای زمانی

مقدمه: اقتصاددانان مشاهده کرده‌اند و معتقدند که مدلهای معادلات همزمان کلان که در سالهای ۱۹۶۰ساخته شده اند، قدرت پیش بینی ضعیف تری ازمدلهای ساده یک متغیره دارند، مدلهایی که صرفاً دارای چند پارامتر و بصورت یک متغیر سری زمانی است.
در روشهای اقتصادسنجی با اسناد کردن متغیرها با یکدیگر، آنها را مورد تجزیه و تحلیل قرار می‌دهند. اسناد کردن متغیرها منبعث از نظریه‌های اقتصادی است. پس ازتخمین پارامتر مدلها، متغیرهای وابسته بر اساس پیش بینی و حرکت متغیرهای مستقل، قابل پیش‌بینی و Forecast می‌باشند. اما هدف پیش بینی را می‌توانیم با یک مدل ساده که رفتار متغیر را بر اساس رفتار آن در گذشته شرح می‌دهد، برآورده نمائیم. بهمین دلیل کارهای (Box-Jenkins, 1984) ، برجستگی خاصی در پیش بینی روند متغیرها پیداکرد. برای رسیدن به تجزیه وتحلیل ماهیت متغیرهای اقتصادی لازم است تعاریف و ابزارهای را در اینجا معرفی نمائیم.

۱- فراگرد تصادفی Stochastic Process

تعریف: مجموعه‌ای ازمتغیرهای تصادفی که دریک فضای احتمالی تعریف می‌شوند را فراگرد تصادفی گویند.

$\{X_t : t \in T\}; T = N$

فضای احتمال (Ω, A, P)

بطوریکه:P مقدار احتمال، A واریانس و Ω فضای نمونه است.

۲- فراگرد تصادفی مانای ضعیفWeakly Stationary Stochastic Process

به یک جریان تصادفی $\{X_t : t \in Z\}$ مانای ضعیف گفته می‌شود اگر:

میانگین آن وجود دارد و به t بستگی ندارد.

$i)\, E(X_t) = \mu$

کوواریانس آن وجود دارد وتابعی ازt-s است.

$$ii)\ \mathrm{var}(X_t-\mu)(X_s-\mu)=\mathrm{Cov}(X_t,X_s)$$

درنتیجه واریانس X_t ثابت است وبه t بستگی ندارد .

$$iii)\ \mathrm{var}(X_t)=E[X_t-\mu][X_t-\mu]=\sigma^2$$

توجه: ازسه شرط فوق نمی توان نتیجه گرفت که گشتاورهای سوم و چهارم هم ثابت هستند. لذا اگـر X_t دارای میانگین μ باشد و X_t ها با یکدیگر رابطه نداشته باشند، و واریانس X_t ثابت باشد، درآنصورت X_t مانای ضعیف است.به ویژه چنانچه اگر X_t , i.i.d با واریانس محدود باشد ، X_t مانای ضعیف است.

۳- فراگرد تصادفی مانای قوی Strictly Stationary Stochastic Process

به یک جریان تصادفی $\{x_t : t\in Z\}$ قویاً مانا گویند، اگـر تـابع چگـالی مشـترک $X_{t1},X_{t2},X_{tk}(cdf)$ برای همه $t_1,t_2,\dots,t_k$ هم مانند تابع چگالی مشترک $x_{t1}+h,x_{t2}+h,x_{tk}+h$ برای مقادیر مختلف h ، باشد.

نکته: هر جریان مانای ضعیف، نمی‌تواند مانای قوی باشد، ولی اگر x_t مانای قـوی باشـد و دارای واریـانس محدود، حتماً مانای ضعیف خواهد بود.

۴- تابع اتوکو واریانس Auto Correlation Function(ACF)

$$E(X_t-\mu)(X_{t+h}-\mu)=\gamma_h$$

بطوریکه $\gamma_{-h}=\gamma_{+h}$

لذا، اگر $h=0$ باشد داریم: $\mathrm{var}(X_t)=\gamma_\circ$

۵- تابع خود همبستگی Auto correlation Function

چنانچه γ_h را بر γ_0 تقسیم نمائیم تابع ACF داریم.

$$\rho_h=\frac{\gamma_n}{\gamma_\circ} \quad ; \quad \rho_h=\rho_{-h} \quad -1\le\rho_h\le 1$$

لذا اگر $h=0$ باشد $p_0=1$ خواهد بود.

۶-تابع خود همبستگی جزیی Partial Autocorrelation Function

تابع ACF(k) ازمرتبه k رابطه ناخالص بین X_t و X_{t-k} رانشان می‌دهد. رابطه ناخالص بـین ایـن دو دوره که همبستگی جزیی نام دارد، عبارتست از همبستگی سادده بین X_t و X_{t-k} منهـای آن بخـش کـه بـا رابطـه خطی بین X_t و X_t های گذشته ، توضیح داده نشده است، یعنی:

$$\rho_h^*=\mathrm{corr}[X_t-\beta_1X_{t-1}-\beta_2X_{t-2}-\dots-\beta_kX_{t-k+1},\quad X_{t-k}]$$

۷- White Noise

اگر فراگرد تصادفی دارای میانگین صفر و واریانس σ^2 باشد و با یکدیگر وابستگی نداشته باشند آنرا White noise گویند و معمولاً آنرا با ε_tیا u_t نشان می‌دهند.

۸- مدلهای اتورگرسیو Autoregressive Model

رابطه زیر راچنانچه ε_t ,white noise باشد برای همه مقادیرt ، یک مدل AR(1) است.

$$X_t = \alpha X_{t-1} + \varepsilon_t$$

مدل بالا اتورگرسیو ازمرتبه اول است.

اگر $|\alpha|<1$ باشد، می‌توان نشان داد که X_t یک جریان تصادفی مانای ضعیف است.

باجایگزاری داریم:

$$X_t = \sum_{j=\circ}^{\infty} \alpha^j \varepsilon_{t-j}$$

بررسی کنیم آیا ویژگیهای مانای ضعیف را دارد؟

$$E(X_t) = \circ \qquad (۱)$$

$$E(X_t X_{t+h}) = E(\sum_{j=\circ}^{\infty} \alpha^j \varepsilon_{t+j} \sum_{k=\circ}^{\infty} \alpha^k \varepsilon_{t+h-k}) \qquad (۲)$$

$$= \sum_{j=\circ}^{\infty} \sum_{k=\circ}^{\infty} \alpha^{j+k} E(\varepsilon_{t-j}\, \varepsilon_{t+h-k})$$

بطوریکه: اگر $t - j = t + h - k$ باشد رابطه فوق σ_ε^2 و اگر $t-j \neq t+h-k$ باشد، برابر صفر خواهد شد.

لذا داریم:

$$= \sum_{j,k=\circ}^{\infty} \alpha^{j+k} \sigma_\varepsilon^2 \;; k = h + j$$

$$= \sigma_\varepsilon^2 \sum_{j,k=\circ}^{\infty} \alpha^{j+k} = \sigma_\varepsilon^2 \sum_{j=\circ}^{\infty} \alpha^{2j+h} = \sigma_\varepsilon^2 \alpha^{|h|} \sum_{j=\circ}^{\infty} \alpha^{2j}$$

$$= \sigma_\varepsilon^2 \alpha^{|h|} \sum_{j=\circ}^{\infty} (\alpha^2)^j = \sigma_\varepsilon^2 \frac{\alpha^{|h|}}{1-\alpha^2} = \gamma_h$$

γ_h اتوکوواریانس X_t مانای ضعیف است.

دقت کنیم که تابع اتوکوواریانس نباید صفر شود، برخلاف تابع خود همبستگی، ACF که ممکن است صفر شود، یعنی ρ_h می‌تواند صفر هم باشد.

می‌دانیم که $\gamma_\circ = \frac{\sigma^2}{1-\alpha^2}$ درنتیجه:

$$\rho_h = \frac{\gamma_n}{\gamma_\circ} = \frac{\sigma^2 \frac{\alpha^{|h|}}{1-\alpha^2}}{\frac{\sigma^2}{1-\alpha^2}} = \alpha^{|h|}$$

لذا $\rho_h = \alpha^{|h|}$ تابعی نزولی از h می‌باشد، چون $|\alpha|<1$ است.

بنابراین درمدل $X_t = \alpha X_{t-1} + \varepsilon_t$ اگر $|\alpha|<1$ باشد، X_t فراگردی ماناست لذا $|\alpha|<1$ شرط پایداری است.

درمراحل بعدی شرط مانایی برای مدلهای اتورگرسیو از مراتب بالاتر راخواهیم گفت. حال اجازه دهید به نوع دیگرِی از فراگردهای تصادفی که به آنها میانگین متحرک می‌گویند، بپردازیم.

۹- فراگرد تصادفی میانگین متحرک Moving Average Stochastic Process

اگر متغیر تصادفی مانند εt ، White noise باشد،

$$E(\varepsilon_t) = 0$$
$$E(\varepsilon_t)^2 = \sigma^2$$
$$E(\varepsilon_t\ \varepsilon_s) = 0 \quad (t \neq s)$$

بطوریکه :

باشد درآنصورت:

فراگرد تصادفی $X_t = \varepsilon_t + \theta\varepsilon_{t-1}$ رافراگرد تصادفی MA(1) ، میانگین متحرک ازمرتبه اول گویند.

ملاحظه می‌شودکه:

$$E(X_t) = 0 \quad (۱)$$

$$E(X_t)^2 = (1+\theta^2)\sigma^2 \quad (۲)$$

$$E(X_t X_{t-1}) = \theta\sigma^2 \quad (۳)$$

بطوریکه: $E(X_t X_{t+2}) = 0$ و $E(X_t X_{t+h}) = 0$ اگر $\Rightarrow h \geq ۲$ یا $h \leq -۲$

پس کوواریانس به t بستگی ندارد ولی به h بستگی دارد، لذا X_t یک جریان ماناست و نشان داده می‌شود، مانای ضعیف است. درمراحل بعد به مراتب بالاتر فراگرد میانگین متحرک می‌پردازیم.

۱۰- فراگرد تصادفی Difference Stationary Process (DSP)

چنانچه سری X_t بصورت زیر باشد:

$$X_t - X_{t-1} = \beta + \varepsilon_t$$

آنرا DSP می‌نامند. ε_t یک سری مانا با میانگین صفر و واریانس σ^2 می‌باشد.

۱۱- گام تصادفی Random walk

حالت خاصی از سری DSP گام تصادفی است. فرض کنیم X_t بصورت $X_t = X_{t-1} + \varepsilon_t$ باشد. یعنی تفاضل اول X_t ماناست. اصطلاحاً این سری را گام تصادفی بدون Drift گویند.

۱۲- فراگرد تصادفی Trend Stationary process

چنانچه سری X_t دارای روند و بصورت زیر باشد:

$$X_t = \alpha + \beta t + \varepsilon_t$$

به آن TSP گویند. ملاحظه می‌شود که هر دو سری فوق دارای Trend[۷۸] می‌باشند، ولی روشی که روند آنها را حذف می‌کنیم متفاوت است. اینکه سری DSP یا TSP باشد دارای کاربردهای اقتصادی است.

قبل از بحث در این مورد بهتر است به سریهای نامانا و ریشه‌های واحد بپردازیم.

۱۳- فراگرد نامانا و ریشه‌های واحد Non- Stationary Process and Unit Roots

فراگرد نامانا و ریشه‌های واحد:

بطور کلی اگر یک فراگرد تصادفی دارای ریشه واحد باشد، یعنی فراگرد اتورگرسیوی است که مقدار جاری یک متغیر از مقدار گذشته آن تبعیت می‌کند.

۱-۱۳- اهمیت ریشه‌های واحد

اهمیت داشتن ریشه واحد برای یک فراگرد تصادفی آنست که اگر شوکی به آن وارد شود اثر دائمی دارد. یعنی اگر متغیر اقتصادی نامانا باشد، دارای ریشه واحد است و شوکهایی که به آن وارد می‌شود دائمی خواهند بود و اگر شوکی دائمی باشد نمی‌تواند ناشی از طرف تقاضا باشد. مثلاً اقتصاددانان معتقدند اگر دولت عرضه پول را ۵ درصد افزایش دهد، قادر نخواهد بود بیکاری را برای همیشه کاهش دهد، در نتیجه پول در بلندمدت مهم نیست. ممکن است دولت قادر باشد در کوتاه‌مدت نوسانات در برخی فعالیت‌های اقتصادی را کاهش دهد، اما در این مدت میزان محصول تغییری نخواهد کرد.

اما Nelson- Plosser (1982) نشان دادند، چون فراگرد تصادفی GNP نامانا است، لذا تغییرات در GNP یا شوک‌هایی که برروی GNP اثر می‌گذارند، شوکهای ناشی از عرضه است. فرض کنیم در زمان معینی مثلاً T شوکی باندازه m (افزایش عرضه پول) برروی آن حادث شود. آنگاه اگر مدل سری GNP بصورت $X_t = X_{t-1} + \varepsilon_t$ باشد، X_t و X_{T+1} و X_{T+2} باندازه m افزایش پیدا می‌کنند، لذا اثر شوک و افزایش عرضه پول در اینجا دائمی است و نظریه "پول در بلندمدت مهم نیست" مورد سئوال واقع می‌شود.

اما اگر سری بصورت مدل: $X_t = \alpha X_{t-1} + \varepsilon_t$ باشد، بطوریکه $|\alpha|<1$ باشد، شوک m در طول زمان از بین خواهد رفت. چون در X_t شوک m وارد می‌شود و در دوره‌های بعدی αm و $\alpha^2 m$ خواهد بود که میرنده است ولذا شوک موقتی خواهد بود می‌گوئیم پول در بلندمدت خنثی است.

لذا چون شوک‌های پولی احتمالاً اثرات دائمی روی GNP ندارد، اگر GNP حقیقی، نامانا و یا DSP باشد، نوسانات در GNP توسط شوک‌های حقیقی باید توضیح داده شود نه توسط شوک‌های اسمی و پولی. پس موضوع اندازه تخمین α اهمیت خاصی پیدا می‌کند و این سئوال مطرح می‌شود که آیا $\alpha = 1$ است ویا کوچکتر از یک

۷۸- غالب متغیرهای اقتصادی عملاً سریهای نامانا هستند، بدین معنی که میانگین و واریانس آنها به زمان بستگی دارد و هر چه زمان پیش می‌رود این متغیرها تمایل دارند که از یک مقدار معینی دور شوند. چنانچه این حرکت در یک جهت (بالا و پایین) باشد، این سری را سری زمانی دارای Trend گویند. لذا برای کارکردن با سریهای زمانی نامانا قبل از هر چیز آنها را de-Trend می‌کنند.

۲-۱۳- مسئله آزمون فرضیه

فرضیه $\alpha=1$ و یا فرضیه مقابل $\alpha<1$، در دو دهه اخیر در اقتصاد کلان اهمیت ویژه‌ای پیدا کـرده اسـت که بیش از ۱۰۰ ها مقاله در این زمینـه تهیـه شـده اسـت و همچنـان مـورد اخـتلاف مکاتـب نئوکلاسـیک، نئوکینزین، Real Business و پول گراها می‌باشد.

برای بررسی این فرضیه دو مشکل آماری وجود دارد:

اول- چگونه روند را حذف کنیم یعنی با تفاضل‌گیری یا با انجام رگرسیون و قـراردادن T بعنـوان متغیـر توضیح دهنده؟

دوم - $\hat{\alpha}_{ols}$ در مدلهای AR چنانچه $\alpha=1$ باشد دارای توزیع استاندارد t و F معمولی نخواهد بود و این توزیع باید مورد به مورد محاسبه شود و بستگی دارد به اینکه سایر متغیرهایی که در مدل وجـود دارنـد کـدام هستند، (آیا جمله ثابت، روند، متغیرهای تاخیری تا چند مرتبه وجود دارند)، لذا نیاز داریم جدول خاصی برای آزمون ریشه واحد تهیه کنیم.

۱۴- آزمون ریشه واحد: Unit Root Test

موضوعی که سالهای ۱۹۸۰ همه انرژی اقتصاددانان را بخود جذب کرد ایـن موضـوع بـود کـه آیـا α در مدلهای $X_t=\alpha X_{t-1}+\varepsilon_t$ برابر یک است یا کوچکتر از یک. میدانیم که اگر $\alpha=1$ باشـد، X_t گـام تصـادفی است و یک حالت خاص از سری نامانا ا ست. همانطور که گفته شد تخمـین α بـه روش OLS دارای توزیـع مجانبی نرمال بصورت زیر است.

$$\sqrt{T}\,(\hat{\alpha}-\alpha) \quad \to \quad N(0,1-\alpha^2)$$

$$P\lim \frac{1}{T}\sum_{t=2}^{T} X_{t-1}^2 = \frac{\sigma^2}{1-\alpha^2}$$

$$P\lim \frac{1}{T}\sum_{t=2}^{T} X_{t-1}\ \varepsilon_t = 0$$

$$P\ \lim\ \hat{\alpha}=\alpha+P\lim\ (\frac{1}{T}\ \Sigma\ {}_{t-1}^{2})^{-1}.P\lim\frac{1}{\sqrt{T}}\ \Sigma X_{t-1}\ \varepsilon_t.P\lim\frac{1}{\sqrt{T}}$$

$$P\ \lim \hat{\alpha}=\alpha$$

اگر سری گام تصادفی باشد، تخمین OLS برای α دارای تـورش بـه سـمت صـفر اسـت. حتـی اگـر α کوچکتر از یک یا نزدیک به یک باشد باز هم تخمین آن به سمت صفر دارای تورش است.

Evans – Said (1981, 1984) با استفاده از روش Monte-Carlo شواهدی نشـان دادنـد کـه $\hat{\alpha}_{ols}$ دارای تورش ا ست. آزمون $H:\alpha=1$ با استفاده از جدول معمولی t و F همواره به نفع فرضیه مقابل رد خواهد شد. Dickey- Fuller (1976) روش آزمونی برای این مسئله ارائه کردند. برای آشنا شدن با این روش مدل:

$$X_t=\beta_\circ+\beta_1 t+u_t$$

$$u_t=\alpha u_{t-1}+\varepsilon_t$$

را درنظر بگیریم. بطوریکه ε_t دارای میانگین صفر و مانا است. فرم خلاصه شده مدل فوق با استفاده ازتبدیل Koick بصورت زیر است.

$$X_t = \gamma + \delta t + \alpha X_{t-1} + \varepsilon_t \quad (۱)$$

$$\gamma = \beta_\circ (1+\alpha)$$

$$\delta = \beta_1 (1-\alpha)$$

مدل فوق دارای ریشه واحد است اگر $\alpha = 1$ باشد. (درآنصورت $\delta = 0$ است). آزمون DF براساس آزمون $\alpha = 1$ در رابطه (۱) می‌باشد تحت این فرض که ε_t , White noise باشد. سه نوع تابع نمونه یا آماره برای این آزمون وجود دارد.

$$k(1) = T(\hat{\alpha} - 1)$$

$$\tau(1) = \frac{\hat{\alpha} - 1}{SE(\hat{\alpha})}$$

$$F(\circ, 1)$$

$\hat{\alpha}$ تخمین ols است و $SE(\hat{\alpha})$ (انحراف معیار $\hat{\alpha}$ می‌باشد.

F ,F(0 ,1) معمولی است برای آزمون مشترک $\begin{cases} \delta = \circ \\ \alpha = 1 \end{cases}$

این آماره‌ها دارای توزیع استاندارد نرمال t و T نیستند. مقادیر بحرانی $k(1)$ و $\tau(1)$ برای $\delta=0$ در مقاله Fuller (1978) جدول بندی شدند و مقادیر بحرانی F (0,1) در مقاله Dickey-Fulle (1981) آمده است.

اما در مدل فوق اولاً ε_t , White noise فرض شده است و ثانیاً ممکن است متغیر با تاخیر از مراتب بالا نیز در سمت راست مدل وجود داشته باشند. بدین منظور،

Philips- Perron(1988), Philips(1987), Dickey- Fuller(1984)

آزمونDF را برای مدلهایی که ضرورتاً ε_t وایت نویز نباشد توسعه دادند و آزمون (Augmented D.F) را ارائه دادند که مدلهای زیر را آزمون می‌کند.

$$X_t = \gamma + \delta t + \alpha X_{t-1} + \sum_{j=1}^{k} \theta_j \Delta X_{t-j} + e_t \quad (۲)$$

هدف از اضافه نمودن جمله ΔX_{t-j} آن است که جریانهای ARMA را مورد بررسی قرار دهیم. اما اگر پارامترهای MA بزرگ باشد، تعریف نمودن AR مدل ضعیفی را ارائه میدهد مگر آنکه k باندازه کافی بزرگ باشد. علیرغم اشکالاتی که به قدرت آزمون ریشه واحد وارد است ، آزمونهای DF و ADF مقبولیت زیادی در بین اقتصاددانان یافته است. قبل از آنکه به معرفی مدلهای ARMA بپردازیم ، اجازه دهید قدرت آزمون ریشه واحد را بررسی کنیم.

۱۵- قدرت آزمون ریشه واحد

در قالب سریهای زمانی متغیرهای اقتصادی ، اقتصاددانان نتوانسته‌اند فرضیه $\alpha = 1$ را رد کنند و لذا نتیجه گرفته اند که تقریباً همه مغیرهای اقتصادی DS می‌باشند. یعنی در سطح ، نامانا و با اولین تفاضل‌گیری مانا می‌شوند.

اما برخی همین نتایج را مربوط نموده‌اند به ضعف آزمون ریشه واحد برای آزمون فرضیه مقابل یعنی TS بودن متغیرها. چون درست است که نمی‌توانند $H_o: \alpha=1$ را رد کنند ولی $\alpha=0.95$ را هم نمی‌توانند رد کنند.

مسئله دیگری که توسط Choi به آن اشاره شده این است که اگر جملات اخلال دارای اجزای MA باشند، یک AR با مراتب بالا مانند آنچه در مدل (۲) می‌بینیم باعث می‌شود که تخمین‌های $\hat{\alpha}_{ols}$ به سمت یک تورش داشته باشد ونشانگر وجود ریشه واحد است حتی اگر $|\alpha|<1$ باشد. Nelson –Plosser (1982) با مدل

$$X_t = \gamma + \delta t + \alpha X_{t-1} + \varepsilon_t \qquad (۳)$$

شروع نمودند و $H_o: \alpha=1, \delta=0$ را تحت شرایط زیر آزمون کردند.

آنها فرض کردند ε_t ,i.i.d با میانگین صفر و واریانس σ^2 باشد و ε_t بصورت:

$$\varepsilon_t = e_t + \phi e_{t-1} \qquad |\phi|<|$$

و این زمانی بود که مشاهده کردند داده‌های آماری ایالات متحده دارای جمله اخلال با تفاضل اول است و سپس بجای مدل (۳) مدل (۲) را به‌منظور جمله اخلال MA مورد استفاده قرار داده‌اند.

بکاربردن مدل (۲) باعث شد که $\hat{\alpha}$ دارای تورش به سمت بالا باشد.

Choi با استفاده از روش Monte Carlo فرض کرد که در مدل (۲):

$$\gamma = 5 \qquad \delta = 1 \qquad \alpha = 0.5$$

و $\phi = 0.3$ دریافتند که $\hat{\alpha}$ در مدل (۲) دارای تورش به سمت یک است.

۱۶- مدلهای ARMA

۱-۱۶- AR(2) و AR(P) : مدل اتورگرسیو مرتبه دوم و مرتبه بالاتر P

موضوع فراگرد تصادفی AR(1) را می‌توان به سادگی به فراگرد AR(2) و یا بیشتر AR(P) تعمیم داد.

فراگرد تصادفی AR(2) را بصورت زیر در نظر بگیریم.

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \varepsilon_t$$

ε_t خصوصیات مطلوب کلاسیک را داراست. با استفاده از عملگر وقفه L رابطه فوق را می‌توان به صورت فشرده زیر نوشت:

$$(1-\alpha_1 L + \alpha_2 L^2) X_t + \varepsilon_t$$

$$A(L) = 1 - \alpha_1 L + \alpha_2 L^2$$

$$A(L)\ X_t = \varepsilon_t$$

از آنجا که A(L) یک معادله درجه ۲ برحسب L است، دارای دو ریشه خواهد بود. یعنی دو جواب برای A(L)=0 وجود خواهد داشت که عبارت است از Z_1 و Z_2 . برای بحث مانایی در این حالت و بررسی پیامدهای آن بهتر است که A(L) را بصورت زیر بنویسیم:

$$A(L) = (1 - r_1 L)(1 - r_2 L)$$

در این صورت جوابهای $A(L)=0$ عبارت خواهد بود از:

$$Z_1=\frac{1}{r_1} \Rightarrow r_1=\frac{1}{Z_1}$$

$$Z_2=\frac{1}{r_2} \Rightarrow r_2=\frac{1}{Z_2}$$

لذا با توجه به روابط بالا می‌توانیم رابطه $AR(2)$ را بصورت زیر بنویسیم:

$$X_t=A(L)^{-1}\varepsilon_t=(1-r_1L)^{-1}(1-r_2L)^{-1}\ \varepsilon_t \qquad (۴)$$

اگر قدر مطلق هر دو مقدار r_1 و r_2 کمتر از یک باشد. وقتی t باندازه کافی بزرگ است با استفاده از تصاعد هندسی می‌توان نوشت:

$$(1-r_1L)^{-1}=\frac{1}{1-r_1L}=1+r_1L+r_1^2L^2+r_1^3L^3+....$$

$$(1-r_2L)^{-1}=\frac{1}{1-r_2L}=1+r_2L+r_2^2L^2+r_2^3L^3+....$$

با جایگذاری روابط فوق در رابطه (۴) خواهیم داشت.

$$X_t=(1+r_1L+r_1^2L^2+r_1^3L^3+....)(1+r_2L+r_2^2L^2+r_2^3L^3+....)\varepsilon_t$$

$$X_t=\delta_0\ \varepsilon_t+\delta_1\ \varepsilon_{t-1}+\delta_2\ \varepsilon_{t-2}+...$$

که در آن داریم:

$$\delta_0=1,\ \delta_1=r_1+r_2\ ,\ \delta_2=t_1^2+r_1\ r_2+r_2^2\ ,\ ...$$

ملاحظه می‌شود که رابطه اخیر فراگرد میانگین متحرک است و هر فراگرد میانگین متحرک ماناست.

بنابراین شرط پایایی برای فراگرد $AR(2)$ آن است که قدر مطلق r_1 و r_2 کوچکتر از واحد باشند. از آنجا که $r_1=\frac{1}{Z_1}$ و $r_2=\frac{1}{Z_2}$ است پس باید قدر مطلق ریشه‌های Z_1 و Z_2 معادله $A(L)=o$ بزرگتر از واحد باشند.

۲-۱۶- مرتبه هم جمعی Integration Order

فرض کنید به این دلیل که یکی از ریشه‌های معادله $A(L)=0$ برابر یک است ، فراگرد $AR(2)$ نامانا است.

بدین ترتیب که $Z_1\ ,\ Z_2\ ,=1$ به صورت قدر مطلق از یک بزرگتر است.

$(|Z_1|>1)$ و یا $(|r_1\ 1<1)$ در چنین حالتی چون $r_2=\frac{1}{Z_2}=1$ است داریم.

$$A(L)X_t=(1-r_1\ L)(1-L)X_t=(1-r_1\ L)\Delta X_t=\varepsilon_t$$

$$A(L)X_t=\Delta X_t-r_1\Delta X_{t-1}=\varepsilon_t$$

$$\Rightarrow \Delta X_t=r_1\Delta X_{t-1}+\varepsilon_t$$

چون $|r_1|<1$ است، پس فراگرد فوق ماناست. بنابراین می‌توان نتیجه گرفت که فراگرد X_t یک فراگرد جمعی از مرتبه یک یا $I(1)$ است و ΔX_t یک فراگرد مانا یا $I(0)$ است.

حال اگر هر دو ریشه معادله $A(L) = 0$ برابر یک باشد، آنگاه $r_1 = r_2 = 1$ خواهد بود. در نتیجه رابطه $A(L)X_t = \varepsilon_t$ به صورت زیر خواهد آمد.

$$A(L)\,X_t = (1-L)\,(1-L) = (1-L)^2\,X_t = \Delta^2 X_t = \varepsilon_t$$

رابطه فوق می‌گوید تفاضل مرتبه دوم X_t فراگرد $I(0)$ و ماناست پس متغیر X_t در سطح فراگرد جمعی از مرتبه ۲ یا $I(2)$ می‌باشد.

لذا مرتبه‌ای که یک فراگرد نامانا را با تفاضل‌گیری مانا کنیم ، مرتبهIntegration یا هم جمعی گویند و بصورت $I(d)$ نشان داده می‌شود.

به سادگی می‌توان فراگرد $AR(p)$ را بسط داد. یعنی اجازه دهیم X_t به وقفه‌های خود تا وقفهPام وابسته باشد. این موضوع به فراگرد خودهمبسته مرتبه Pام زیر می‌انجامد.

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \ldots + \alpha_p X_{t-p} + \varepsilon_t$$

۳-۱۶- فراگرد میانگین متحرک مرتبه q ام، MA(q)

وقتی که X_t به صورت تابعی از وقفه‌های جملات اخلال همبسته باشد، X_t یک فراگرد میانگین متحرک است وچنانچه تعداد این وقفه‌ها q باشد آنرا $MA(q)$ گویند.

$$X_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \ldots + \theta_q \varepsilon_{t-q}$$

که در آن ε_t ها ناهمبسته با میانگین صفر و واریانس σ^2 هستند. تمامی فرایندهای میانگین متحرک مانا هستند.

۴-۱۶- مدلهای Autoregressive Moving Average Process

چنانچه دو فراگرد تصادفی AR و MA با ویژگیهای مختلف را باهم تلفیق نمائیم ، تشکیل یک فراگرد ARMA را میدهد. به عنوان مثال ARMA(۲، ۱) به صورت زیر نوشته می‌شود.

$$X_t = \alpha X_t + \varepsilon_{t-1} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2}$$

فراگرد تصادفی(ARMA) از انعطاف پذیرترین نوع الگوهای سریهای زمانی تک متغیره است و یک سری $ARMA\ (P, q)$ بصورت زیر است.

$$(۵)\quad X_t = \alpha_1 X_{t-1} + \ldots + \alpha_p X_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q}$$

مشروط بر اینکه پارامترهای $\theta_1 \ldots \theta_q$ شرایط خاصی را تامین کنند. می‌توان فراگرد تصادفی (۵) را به یک فراگرد تصادفی اتورگرسیو تبدیل کرد.

۵-۱۶- فراگرد تصادفیARIMA

چنانچه یک سری زمانی را dبار تفاضل‌گیری کنیم ، یعنی هم جمع از مرتبه d باشد و آنگاه آن را در قالب مدل $ARMA\ (p\ ,\ q)$ بیآوریم، می‌گوئیم که سری زمانی اولیه یک فراگرد اتورگرسیو میانگین متحرک از مرتبه d می‌باشد.

که با مرتبه اتورگرسیو P و میانگین متحرک q آنرا بصورت $ARIMA\ (p\ ,\ d\ ,\ q)$ نمایش می‌دهند.

۶-۱۶- فرآگرد تصادفی اتورگرسیو برداری Vector Auto regressive Process (VAR)

مطالب ارائه شده در مورد فراگردهای اتورگرسیو و میانگین متحرک و همچنین مدلهای ادغامی این دو را می‌توان به سریهای زمانی چند متغیره نیز بسط داد.

چنانچه X_t برداری از متغیرهای سری زمانی باشد که بتوان آنرا توسط وقفه‌های آن بصورت یک فراگرد چند متغیره (یا برداری) ARMA مدل‌سازی کرد، در آنصورت می‌توان نوشت:

$$X_t = \Phi_1 X_{t-1} + \dots + \Phi_p X_{t-p} + \varepsilon_t + \Theta_1 \varepsilon_{t-1} + \dots + \Theta_q \varepsilon_{t-q} \quad (۶)$$

که در آن Φ_1 , Φ_2 , … , Φ_p , Θ_1 , Θ_2 , … , Θ_q ماتریسهای مربع شامل پارامترها و ε_t بردار مربوط به جملات اخلال است. یک مدل VAR را معمولاً بصورت زیر می‌نویسند:

$$X_t = \Phi_1 X_{t-1} + \dots + \Phi_3 X_{t-p} + \varepsilon_t \quad (۷)$$

معمولاً برآورد چنین الگوهایی نسبتاً مشکل است.

رگرسیون هم جمعی The Co-Integration Regression

برای گرفتن ایده کلی مثال ساده‌ای که Engale-Granger (1987) ارائه کردند را اریه میکنیم:

فرض کنیم x_{t1} , x_{t2} دو سری باشند که بصورت زیر ارائه شده باشند ودو جمله وایت نویز e_{t1} , e_{t2} را که ممکن است با یکدیگر رابطه داشته باشند را در نظر بگیریم:

.

$$x_{t_1} + \beta x_{t_2} = u_{t_1} \qquad u_{t1} = u_{t-1,1} + e_{t1} \quad (۱)$$

$$x_{t_1} + \alpha x_{t_2} = u_{t_2} \qquad u_{t2} = \rho u_{t-2,1} + e_{t2} \quad |\rho| < | \quad (۲)$$

بطوریکه u_t دارای مرتبه همجمعی یک $u_{t1} \sim I(1)$ و $u_{t2} \sim I(0)$ میباشند.

مدل فوق سازگار خواهد بود اگر $\alpha \neq \beta$ باشد. این قید برای این است که اگر $\alpha = \beta$ باشد، غیرممکن است مقادیری برای x_{t_1} و x_{t_2} پیدا کرد که بطور همزمان در دو تساوی فوق صدق نمایند.

به مفهوم عام β و α به دلیل عدم وجود متغیر برونزا در مدل غیرقابل شناسایی هستند.

فرم خلاصه شده مدل فوق بصورت زیر است.

$$x_{t1} = \frac{\alpha}{\alpha - \beta} u_{t1} - \frac{\beta}{\alpha - \beta} u_{t2} \quad (۱)$$

$$x_{t2} = \frac{1}{\alpha - \beta} u_{t1} + \frac{1}{\alpha - \beta} u_{t2} \quad (۲)$$

دو متغیر فوق ترکیب خطی از u_{t2} و u_{t1} می‌باشند و در نتیجه هر دو $I(1)$.

توجه کنیم که معادله ۲ نشانگر ترکیب خطی دو متغیری است که هر دو $I(1)$ و نامانا هستند لذا x_{t_1} و x_{t_2} با یکدیگرهمجمع (Co-Integrate) می‌باشند.

در این شرایط رگرسیون x_{t_1} برروی x_{t_2} به روش حداقل مربعات تخمینی از α بدست می‌دهد که آنرا Supper-Constient می‌گویند یعنی اینکه سریعتر از $\hat{\alpha}_{ols}$ تمایل دارد که به مقدار واقعی α میل کند.

در موارد معمولی، اگر $\hat{\beta}$ تخمین حداقل مربعات از β باشند.

$\sqrt{T}\ (\hat{\beta}-\beta) \to 0$ وقتی $T \to \infty$

درصورتیکه در اینجا $T\ (\hat{\beta}-\beta) \to 0$ وقتی $T \to \infty$

رگرسیون x_{t_1} برروی x_{t_2} را در اینجا Co-integration-Regression می‌نامند.

تمامی ترکیبات خطی x_{t_1} و x_{t_2} غیر از مدل (۲) دارای واریانس نامحدودی هستند.

توجه داریم که اگر $\rho=1$ باشد u_{t_1} نیز $I(1)$ است و در نتیجه Co-integration-Regression نخواهیم داشت. اما می‌توانیم معادلات (۱) و (۲) را بصورت مدلهای AR بنویسیم.

$$\Delta x_{t_1} = \beta\delta x_{t-1,1} + \alpha\ \beta\delta x_{t-1,2} + \eta_{t_1} \qquad (۳)$$

$$\Delta x_{t_2} = -\delta x_{t-1,1} - \alpha\ \delta x_{t-1,2} + \eta_{t_2} \qquad (۴)$$

بطوریکه: $\delta = (1-\rho)/\alpha - \beta$

η_{t_2} و η_{t_1} ترکیبات خطی e_t می‌باشد. معادلات (۳) و (۴) نشان‌دهنده یک مدل VAR برای این روابط ساده است اگر چنانچه $z_t = z_{t_1} + \alpha x_{t_2}$ تعریف کنیم می‌توانیم داشته باشیم:

$$\Delta x_{t_1} = \beta\delta z_t + \eta_{t_1} \qquad (۵)$$

$$\Delta x_{t_2} = \beta\delta z_{t-1} + \eta_{t_2} \qquad (۶)$$

معادلات (۵) و (۶) نشانگر ECM (Error Correction Model) مدل فوق هستند. الا اینکه در این مدل Δx_{t_1} با Δx_{t_2} درگیر نیست و بالعکس آن نیز صحیح است.

اگر مدل فوق را با OLS تخمین بزنیم چون η_{t_2} و η_{t_1} نامرتبط هستند. برآورد پارامترها سازگار خواهند بود. لذا تخمین سازگاری از β در معادله (۵) می‌توانیم بدست آوریم.

یک مدل ECM بصورت $\Delta y_t = \gamma\Delta x_t + \alpha(y-\beta x_t) + u_t$ رابطه‌ای بین تغییرات y_t و تغییرات x_t و عدم تعادل در گذشته است. ECM فوق بصورت زیر استخراج شده است.

$$z_t = x_{t_1} + \alpha x_{t_2}$$

لذا با استفاده از معادله (۲) داریم:

$$\Delta z_t = (\rho-1)\ z_{t-1} + e_{t_2}$$

یا:

$$\Delta x_{t_1} = \alpha\Delta x_{t21} + (\rho-1)\ z_{t-1} + e_{t_2}$$

این معادله را اگر با ols تخمین بزنیم پارامترها، سازگار نخواهند بود. چون e_{t2} و x_{t2} با هم مرتبط هستند.

ازطرفی متغیرها در این مدل همگی $I(0)$ هستند.

یادآوری (۲): اگر x_t و y_t ، هم جمع باشند درآن صورت رابطه بلند مدتی بین آنها برقرار است. بعلاوه رابطه کوتاه مدت و پویایی از طریق مدل ECM می‌توان برای آنها بیان نمود.

این موضوع به Granger Representation Theorem شناخته شده است.

اگر

$$[z_t = y_t - \beta x_t] \sim I(0)$$
$$y_t \sim I(1)$$
$$x_t \sim I(1)$$

درآنصورت y , x با یکدیگرهمجمع میشوند.

قضیه GRT میگوید که در این مورد x_t و y_t می‌توانند بصورت ECM از فرم زیر در نظر گرفته شوند.

$$\Delta x_{t1} = P_1 z_{t-1} + \text{lagged}\,(\Delta x_t, \Delta y_t) + \varepsilon_{t1}$$
$$\Delta y_t = P_2 z_{t-1} + \text{lagged}\,(\Delta x_t, \Delta y_t) + \varepsilon_{t2}$$

بطوریکه P_1 و P_2 غیر صفر می‌باشند، ε_{t2} و ε_{t1} جملات خطای وایت نویز می‌باشند.

یادآوری (۳): چگونه معادله (۱) و(۲) را شناسایی کنیم یعنی α و β چگونه شناسایی میشوند.

جواب این است که ما با استفاده از اطلاعاتی که درباره جملات اضافی داریم این کار را انجام میدهیم.

چون u_{t1} ، گام تصادفی و u_{t2} , $I(0)$ است.

هر چند با در نظر گرفتن ترکیب خطی دو معادله می‌توانیم معادله‌ای بوجود آوریم که شبیه به یکدیگر هستند. ولی هیچ ترکیب خطی نمی‌تواند جمله خطای $I(0)$ شبیه آنچه در معادله دوم وجود دارد تولید کند. لذا α قابل شناسایی است.

بدین ترتیب هیچ ترکیب خطی نمی‌تواند جمله خطای $I(1)$ یعنی Random walk شبیه آنچه در معادله اول می‌باشد، بوجود آورد لذا β نیز قابل شناسایی است.

معادله (۲) با OLS قابل تشخیص است و $\hat{\alpha}$ سازگار خواهد داد و بدلیل ماهیت x_{t2} که $I(1)$ است و u_{t2} که $I(0)$ است مشکل تورش همزمانی نخواهیم داشت.

آنگاه $z_t = x_{t1} + \alpha x_{t2}$ را تشکیل می‌دهیم و تخمین β را از معادله (۳) بدست می‌آوریم.

Engale-Granger پیشنهاد می‌کنند که برای تخمین ابتدا Co-integration Regression را تخمین بزنیم سپس مدل کوتاه مدت پویا را از طریق تغییرات یک مدل ECM توسط یک روش دو مرحله‌ای که در مرحله اول از برآوردهای CO-R بکار ببریم ، استفاده کنیم.

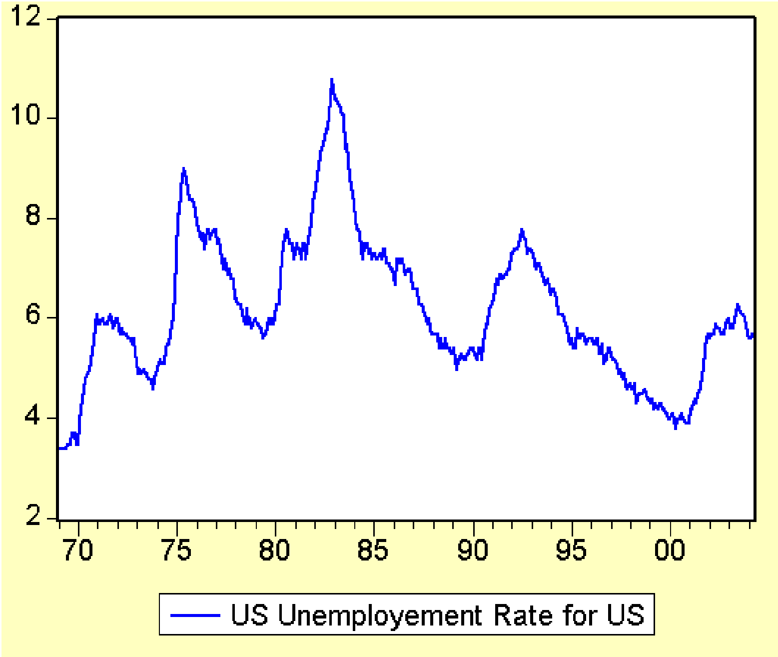
دیگران پیشنهاد می‌کنند که پارامترهای بلند مدت و پویای کوتاه مدت را بطور همزمان تخمین بزنیم.

مثلاً Banrjee (1986) و دیگران با استفاده از مطالعات M.C معادلاتی شبیه به (۱) و(۲) را بررسی نمودند و به این نتیجه رسیدند که در نمونه‌های کوچک تخمین α از معادله (۲) دارای تورش است و پیشنهاد نمودند که بهتر است پارامتر های بلند مدت را از طریق یک مدل پویا تخمین بزنیم.

مثال ۱

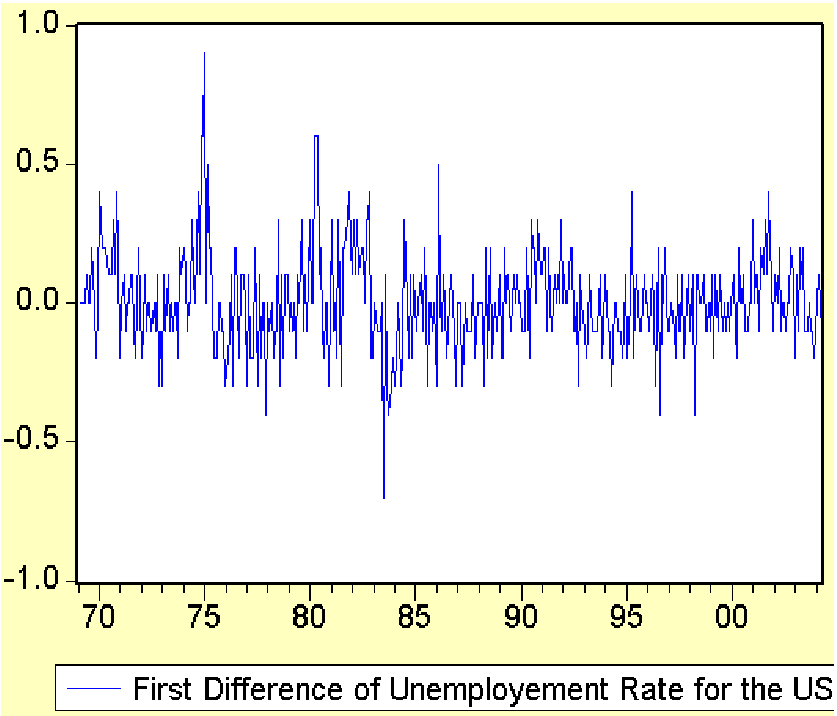
شکل ۱

در این شکل ترسیم متغیر نرخ بیکاری برای ایالات متحده از سال ۱۹۶۹ فصل اول تا فصل چهارم سال ۲۰۰۴ نشان داده شده است. ملاحظه میشود که این متغیر در طول زمان تمایلی به حرکت پیرامون میانگینش ندارد. پس متغیربیکاری در سطح **نامانا** است.



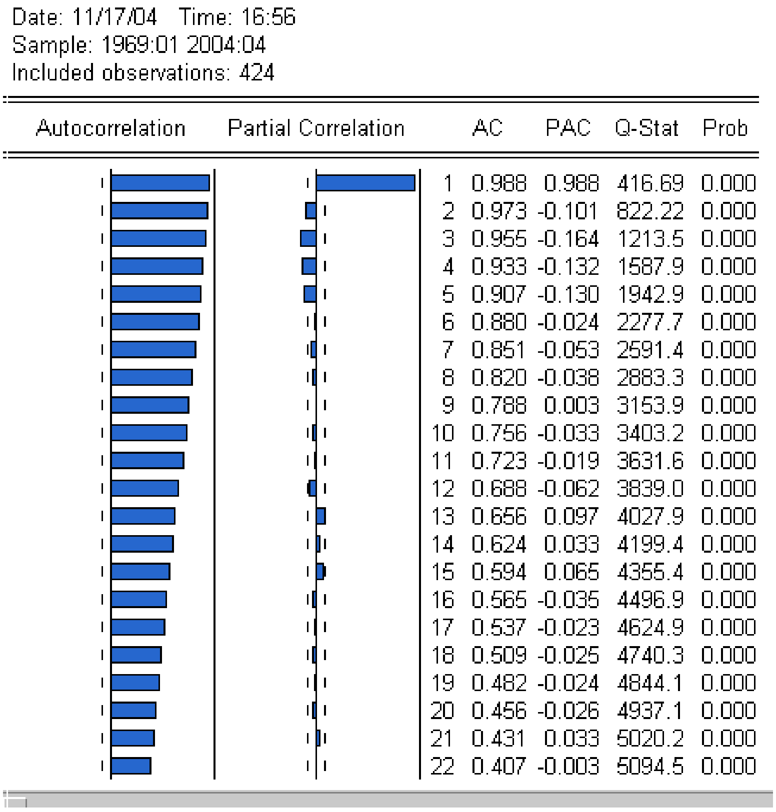
شکل ۲

در این شکل ترسیم متغیر تفاضل اول نرخ بیکاری برای ایالات متحده از سال ۱۹۶۹ فصل اول تا فصل چهارم سال ۲۰۰۴ نشان داده شده است. ملاحظه میشود که این متغیر در طول زمان دارای حرکتی پیرامون میانگینش دارد یعنی متغیر نرخ بیکاری بیکاری با اولین تفاضل گیری **مانا** میشود.



شکل ۳

شکل **Correlogram** برای تشخیص مانایی متغیر نرخ بیکاری در این شکل ترسیم شده است. ملاحظه فرم دنده اره ای برای **ACF** نشانگر **نامانا** بودن این متغیر است. **ACF** ها حتی تا ۲۴ تاخیر به صفر نزدیک نمی شوند.



Date: 11/17/04 Time: 16:56
Sample: 1969:01 2004:04
Included observations: 424

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.988	0.988	416.69	0.000
		2	0.973	-0.101	822.22	0.000
		3	0.955	-0.164	1213.5	0.000
		4	0.933	-0.132	1587.9	0.000
		5	0.907	-0.130	1942.9	0.000
		6	0.880	-0.024	2277.7	0.000
		7	0.851	-0.053	2591.4	0.000
		8	0.820	-0.038	2883.3	0.000
		9	0.788	0.003	3153.9	0.000
		10	0.756	-0.033	3403.2	0.000
		11	0.723	-0.019	3631.6	0.000
		12	0.688	-0.062	3839.0	0.000
		13	0.656	0.097	4027.9	0.000
		14	0.624	0.033	4199.4	0.000
		15	0.594	0.065	4355.4	0.000
		16	0.565	-0.035	4496.9	0.000
		17	0.537	-0.023	4624.9	0.000
		18	0.509	-0.025	4740.3	0.000
		19	0.482	-0.024	4844.1	0.000
		20	0.456	-0.026	4937.1	0.000
		21	0.431	0.033	5020.2	0.000
		22	0.407	-0.003	5094.5	0.000

شکل ۴

در این شکل **AFC** ها با شتاب بیشتری به سمت صفر شدن تمایل دارند. این فرم**Correlogram** تفاضل اول متغیر نرخ بیکاری است که به صورت **مانا** در آمده است.

Date: 11/17/04 Time: 17:06
Sample: 1969:01 2004:04
Included observations: 423

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.154	0.154	10.096	0.001
		2	0.254	0.236	37.595	0.000
		3	0.228	0.176	59.863	0.000
		4	0.248	0.167	86.221	0.000
		5	0.126	0.007	93.050	0.000
		6	0.148	0.021	102.49	0.000
		7	0.108	-0.002	107.56	0.000
		8	0.078	-0.023	110.17	0.000
		9	0.075	0.004	112.59	0.000
		10	0.004	-0.065	112.59	0.000
		11	0.083	0.049	115.63	0.000
		12	-0.071	-0.106	117.80	0.000
		13	-0.001	-0.022	117.80	0.000
		14	-0.097	-0.091	121.98	0.000
		15	0.022	0.051	122.18	0.000
		16	-0.020	0.052	122.36	0.000
		17	-0.019	0.012	122.52	0.000
		18	-0.034	0.000	123.03	0.000
		19	-0.002	0.005	123.04	0.000
		20	-0.041	-0.030	123.77	0.000
		21	-0.026	-0.006	124.07	0.000
		22	-0.012	-0.005	124.13	0.000

جدول ۵

روش آزمون ریشه واحد برای بررسی مانایی متغیر نرخ بیکاری .

با توجه به اینکه قدر مطلق ADF از مقادیر بحرانی ADF کوچکتر است .

دال بر وجود ریشه واحد است. و این تایید دیگری است بر نامانا بودن نرخ بیکاری.

ADF Test Statistic	-3.542231	1% Critical Value*	-3.9842
		5% Critical Value	-3.4224
		10% Critical Value	-3.1338

*MacKinnon critical values for rejection of hypothesis of a unit root.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(UNEMR)
Method: Least Squares
Date: 11/17/04 Time: 18:05
Sample(adjusted): 1969:09 2004:04
Included observations: 416 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
UNEMR(-1)	-0.022015	0.006215	-3.542231	0.0004
D(UNEMR(-1))	0.036696	0.048931	0.749938	0.4537
D(UNEMR(-2))	0.167177	0.049016	3.410687	0.0007
D(UNEMR(-3))	0.155683	0.049710	3.131822	0.0019
D(UNEMR(-4))	0.168565	0.049715	3.390619	0.0008
D(UNEMR(-5))	0.021160	0.049948	0.423647	0.6720
D(UNEMR(-6))	0.039231	0.049346	0.795026	0.4271
D(UNEMR(-7))	0.018526	0.049369	0.375250	0.7077
C	0.165105	0.046161	3.576724	0.0004
@TREND(1969:01)	-0.000122	7.19E-05	-1.692010	0.0914

R-squared	0.158918	Mean dependent var	0.005048
Adjusted R-squared	0.140273	S.D. dependent var	0.181157
S.E. of regression	0.167971	Akaike info criterion	-0.706302
Sum squared resid	11.45503	Schwarz criterion	-0.609411
Log likelihood	156.9108	F-statistic	8.523494
Durbin-Watson stat	1.997169	Prob(F-statistic)	0.000000

شکل ۶

روش آزمون ریشه واحد برای بررسی مانایی تفاضل اول متغیر نرخ بیکاری .
با توجه به اینکه قدر مطلق ADF در سطوح مختلف مقادیر بحرانی بزرگتراست .
دال بر عدم وجود ریشه واحد است. و این تایید دیگری است که نرخ بیکاری پس از یک بار تفاضل گیری مانا میشود.

ADF Test Statistic	-5.237386	1% Critical Value*	-3.9842
		5% Critical Value	-3.4224
		10% Critical Value	-3.1338

*MacKinnon critical values for rejection of hypothesis of a unit root.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(UNEMR,2)
Method: Least Squares
Date: 11/17/04 Time: 18:16
Sample(adjusted): 1969:10 2004:04
Included observations: 415 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(UNEMR(-1))	-0.449947	0.085911	-5.237386	0.0000
D(UNEMR(-1),2)	-0.505182	0.087601	-5.766853	0.0000
D(UNEMR(-2),2)	-0.331242	0.088691	-3.734789	0.0002
D(UNEMR(-3),2)	-0.175370	0.087528	-2.003588	0.0458
D(UNEMR(-4),2)	-0.009561	0.085023	-0.112458	0.9105
D(UNEMR(-5),2)	0.000512	0.079157	0.006468	0.9948
D(UNEMR(-6),2)	0.024995	0.068887	0.362841	0.7169
D(UNEMR(-7),2)	0.023949	0.049759	0.481297	0.6306
C	0.011765	0.017440	0.674573	0.5003
@TREND(1969:01)	-4.66E-05	7.05E-05	-0.660885	0.5091

R-squared	0.487921	Mean dependent var	-0.000723
Adjusted R-squared	0.476541	S.D. dependent var	0.235650
S.E. of regression	0.170494	Akaike info criterion	-0.676434
Sum squared resid	11.77260	Schwarz criterion	-0.579367
Log likelihood	150.3601	F-statistic	42.87698
Durbin-Watson stat	1.998479	Prob(F-statistic)	0.000000

مثال ۲

دیکی-فولر لگاریتم متغیر **Quarterly Federal Reserve Board Production Index (QFRBPI)** را برای فصول اول تا چهارم سالهای 1977 – 1950 بکار بردند و فرض نمودند که این متغیر بصورت مدل زیر است.

$$y_t = \beta_\circ + \beta_1 t + \alpha_1 y_{t-1} + \alpha_2 (y_{t-1} - y_{t-2}) + e_t \qquad AR(2) \qquad e_t \sim IN(0, \sigma^2)$$

تحت شرایط H_1 تخمین ols پارامترهای فوق بصورت زیر میباشد.

$$y_t = 0.52 + 0.00120t - 0.119\, y_{t-1} + 0.498\, (y_{t-1} - y_{t-2})$$

$$(0.15) \quad (0.00034) \quad (0.033) \quad (0.081)$$

$RSS = 0.056448$ مجوع مجذورات خطا

تحت شرایط H_0 داریم:

$$y_t - y_{t-1} = 0.0054 + 0.447\left(y_{t-1} - y_{t-2}\right)$$
$$(0.0025) \quad (0.083)$$

$$RSS = 0.63211$$

$$F = \frac{0.63211 - \frac{0.056448}{2}}{\frac{0.05448}{106}} = 6.34$$

از جدول معمولی F برای درجه آزادی (2, 106) در سطح ۱٪ جدول F > F محاسبه شده است. اما اگر از جدول D.F استفاده شود. F = 6.49 در سطح ۵٪ لذا فرضیه اینکه AR(2) فوق دارای ریشه واحد است در سطح ۵٪ پذیرفته شده است و در سطح ۱۰٪ رد می‌شود.

پیوست شماره ۱

مروری بر جبر ماتریس‌ها

۱-۱ مفاهیم و نمادهای ابتدایی

تعریف (۱): ماتریس – آرایه مستطیل وار اعداد (حقیقی) m در n با نماد α_{ij}، ماتریس حقیقی $n \times m$ بعدی گفته می‌شود. a_{ij} ها، درایه‌های ماتریس گفته می‌شوند.

ستون j به این صورت نشان داده می‌شود:

$$\alpha_{.j} = \begin{bmatrix} a_{ij} \\ \vdots \\ a_{mj} \end{bmatrix}$$

و ردیف i به این صورت نشان داده می‌شود:

$$a_{i.} = [a_{i1} \dots a_{in}]$$

تعریف (۲): فرض کنید A یک ماتریس حقیقی$m \times n$ است.

الف) اسکالر- درایه‌های یک ماتریس حقیقی اسکالر گفته می‌شود.

ب) بردار- به ماتریس‌هایی که یا یک ردیف هستند و یا یک ستون، بردار می‌گویند، یعنی یا $m = 1$ یا $n = 1$

ب) هم ارزی ماتریسها- ماتریس‌های A و B هم ارزنده هر گاه دو ماتریس هم بعد باشند و:

$$\forall_{ij} \Rightarrow a_{ij} = b_{ij}$$

ت) ترانهاده – ترانهاده ماتریس$A(m \times n)$ ،که به صورت A' یا A^T نشان داده شده است، به این صورت تعریف می‌شود:

$$A'=\begin{bmatrix} a_{11} & \cdots & a_{m1} \\ \vdots & \ddots & \vdots \\ a_{1n} & \cdots & a_{mn} \end{bmatrix}$$

توجه داریم که A'، $n \times m$ است.

ث) ماتریس مربع- اگر m = n باشد، ماتریس مربع است.

ج) ماتریس متقارن- ماتریس مربعی که در آن $A=A'$ باشد.

چ) ماتریس قطری- این نوع ماتریس، ماتریس مربعی است که همه درایه‌های غیرقطری آن صفر باشد، یعنی برای هر $i \neq j$، $a_{ij} = 0$. سپس می‌نویسیم: $A=diag\,(a_{11}\ldots,a_{mm})$

ح) ماتریس بالا مثلثی و پایین مثلثی – ماتریس مربعی را که برای هر $i>j$ داشته باشیم $a_{ij} = 0$، بالا مثلثی و ماتریس مربعی را که برای هر $i<j$ داشته باشیم $a_{ij} = 0$، پایین مثلثی می‌گوییم.

خ) ماتریس یکه – ماتریس مربعی را که به صورت زیر باشد، ماتریس یکه از مرتبه n می‌گوییم.

$$I_n = \begin{bmatrix} 1 & o & \cdots & o & o \\ o & 1 & & & \\ \vdots & \vdots & \ddots & & \\ o & o & & 1 & o \\ o & o & \cdots & o & 1 \end{bmatrix}_{n\times n}$$

د) ماتریس تهی - به ماتریس $n \times m$ که به صورت زیر باشد، ماتریس $n \times m$ تهی (صفر) می‌گوییم.

$$o=\begin{bmatrix} o & \cdots & o \\ \vdots & \ddots & \vdots \\ o & \cdots & o \end{bmatrix}_{m\times n}$$

۲-۱. خلاصه‌ای از عملیات پایه‌ای ماتریس

تعریف (۱): ضرب اسکالر- اگر A یک ماتریس حقیقی $n \times m$ باشد و C یک عدد حقیقی، خواهیم داشت:

$$CA=[ca_{ij}]\mathrm{m}\times \mathrm{n}$$

تعریف (۲): جمع ماتریس‌ها- اگر A و B ماتریس‌های حقیقی $n \times m$ باشند، خواهیم داشت:

$$A+B=[a_{ij}+b_{ij}]$$

تعریف (۳): ضرب ماتریس – اگر A و B ماتریس‌های حقیقی $n \times m$ و $q \times n$ باشند، خواهیم داشت:

$$AB=\left[\sum_{k=1}^{n} a_{ik}\, b_{kj}\right]_{m\times q}$$

قضیه(۱): خواص جمع و ضرب – اگر ماتریس‌های C, A, B در ابعاد مناسب باشند، داریم:

قانون شرکت پذیری جمع $\quad I.(A+B)+C=A+(B+C)$

قانون تعویض پذیری جمع $II. A+B=B+A$

قانون شرکت پذیری ضرب $III. (AB)\,C = A\,(BC)$

قانون توزیع پذیری $IV. A(B+C)=AB+AC$

$V.\ (A+B)C=AC+BC$

$VI. A+O=A$

$VII.\ AO=OA=O$

نکته ← قانون تعویض پذیری در مورد ضرب ماتریس‌ها صدق نمی‌کند، در واقع $AB\neq BA$

قضیه (۲). خواص ترانهاده – اگر ماتریس‌های C, B, A در ابعاد مناسب باشند، داریم:

$I.\ C=AB \Rightarrow C'=B'A'$

$II. C=A+B \Rightarrow C'=A'+B'$

$III. C=A' \Rightarrow C'=A$

$IV. (C')'=C$

۳-۱. وابستگی خطی بردارها و مرتبه ماتریس

دستگاه m معادله‌ای ناهمگن زیر را در نظر بگیرید:

$$a_{11}X_1 + \ldots + a_{1n}X_n = 0\,,$$
$$\vdots$$
$$a_{m1}X_1 + \cdots + a_{mn}X_n = 0,$$

یا در قالب ماتریس می‌نویسیم:

$$AX=0$$

که در اینجا $X=[X_1,\cdots,X_n]'$ و 0 برداری $1\times m$ از صفر است یا:

$$a_1X_1+\cdots+a_nX_n=0$$

تعریف (۱): ستون‌های مستقل خطی – اگر هیچ بردار $X\neq 0$ وجود نداشته باشد (یعنی برداری که حداقل یک $X_i\neq 0$ داشته باشد) به طوری که $AX=0$ باشد، می‌گوییم ستونهای A از نظر خطی مستقل‌اند؛ در غیر این صورت ستونها وابستگی خطی دارند.

حال، دستگاه m معادله‌ای همگن زیر را در نظربگیرید:

$$a_{11}y_1+\cdots+a_{m1}y_m=0$$
$$\vdots$$
$$a_{1n}y_1+\cdots+a_{mn}y_m=0$$

یا در قالب ماتریسی می‌نویسم:

$$A'Y=0$$

که در اینجا $Y=[y_1\cdots,y_m]'$ و 0 یک بردار $1\times n$ از صفر است یا:

$$a_1'y_1+\cdots+a_m'y_m=0$$

تعریف (۲): ردیفهای مستقل خطی – اگر هیچ بردار $y\neq 0$ وجود نداشته باشد (مثلا برداری که حداقل یک $y_i \neq 0$ داشته باشد) به طوری که $A'Y=0$ باشد، می‌گوییم ردیفهای A از نظر خطی مستقل هستند؛ در غیر این صورت ردیفها وابستگی خطی دارند.

تعریف (۳): رتبه ردیف و ستون – رتبه ستون ماتریس $A(m\times n)$، مساوی است با حداکثر تعداد ستونهای مستقل خطی. رتبه ردیف ماتریس $A(m\times n)$ مساوی است با حداکثر تعداد ردیفهای مستقل خطی.

قضیه (۱): برای هر ماتریس، حداکثر تعداد ستونهای مستقل خطی مساوی است با حداکثر تعداد ردیف مستقل خطی، مثلا رتبه ستونی A برابر است با رتبه ردیفی A.

قضیه بالا تعریف بعدی را مجاز می‌شمارد.

تعریف (۴): رتبه یک ماتریس – رتبه ماتریس A، یا r(A)، حداکثر تعداد ردیفها یا ستونهای مستقل خطی تعریف می‌شود.

رتبه A = r (A) = رتبه ستون A = رتبه ردیف A

توجه : از تعویض پذیری ستونها و ردیف‌ها در مشخص کردن رتبه، روشن می‌شود که اگر $A\times m$,n باشد، داریم:

$$r(A)\leq \min\ (m\ ,n).$$

تعریف (۵): رتبه کامل - اگر A یک ماتریس m × n و $m\leq n$ باشد، می‌گوییم ماتریس A دارای رتبه کامل است، اگر $r(A)=m$ باشد (پس یک ماتریس مربع m ×m دارای رتبه کامل است، اگر r (A) = m باشد).

تعریف (۶): ماتریس غیر منفرد - اگر A یک ماتریس m ×m باشد، A را غیر منفرد می‌خوانیم، اگر r(A) = m باشد.

قضیه (۲): رتبه حاصل ضرب ماتریس– اگر ماتریسهای C,B, A دارای ابعاد مناسب باشند، داریم:

$$r(AB)\leq \min\{r(A),r(B)\}$$

چنانچه B غیرمنفرد باشد، در نتیجه $r(A)\ =r(AB)$

چنانچه C غیر منفرد باشد، در نتیجه $r(A)=r(CA)$

چنانچه C، B غیرمنفرد باشد، در نتیجه $r(A)=r(CAB)$

$$r(A'A)=r(A)=r(A')$$

تعریف (۷): فضای هیچ و هیچه– اگر ماتریس $A\times p$, m باشد، فضای هیچ A، بگوییم N(A)، به عنوان مجموعه بردارهای حاصل{x : AX = 0} تعریف شده است؛ هیچه A، بگوییم N(A) ، به عنوان حداکثر تعداد بردارهای خطی مستقل در N(A) تعریف شده است. حال می‌توانیم نتیجه زیر را در مورد حاصل دستگاه های معادلات همگن بکاربگیریم.

قضیه (۳): اگر A یک ماتریس p×m باشد، در نتیجه r (A) + n (A)= m خواهد بود. بنابر این، دستگاه معادلات همگن AX = 0 (با فرض سازگار بودن)، یک راه حل منحصر به فرد (غیراز اسکالر) دارد، اگر و فقط اگر r(A) = m-1 باشد.

۴-۱- دترمینان ماتریس

تعریف (۱): دترمینان- اگر A یک ماتریس m ×m باشد، دترمینان ماتریس A، که به صورت $|A|$ یا (A) det نوشته میشود، به این صورت تعریف می‌گردد:

$$|A| = \Sigma (-1)^s a_{1j_1} \dots a_{mj_m}$$

که در اینجا $j_1 \dots j_m$ یک آرایش (جایگشت) ازشماره‌های m، ... ، ۱،۲ است و S صفر یا یک است، بسته به این که عدد ترانهاده لازم برای اعاده دادن $j_1, \dots, j_m$ به ترتیب طبیعی m، ...، ۳،۲،۱ زوج یا فرد باشد؛ مجموع آن بر تمام آرایش‌های ممکن گرفته می‌شود (اشاره به این نکته الزامی است که دترمینان یک ماتریس حقیقی، یک عدد حقیقی است).

مثال: در صورت داشتن یک ماتریس دو یا سه بعدی، داریم:

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \Rightarrow |A| = a_{11}a_{22} - a_{12}a_{21}$$

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \Rightarrow |A| = a_{11}a_{22}a_{33} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31}$$

$$- a_{11}a_{22}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32}$$

برای محاسبه دترمینان ماتریس، می‌توانیم از روش بسط هم عامل‌ها که به صورت زیر توصیف می‌شود، استفاده کنیم:

تعریف (۲): کهاد و هم عامل - اگر A یک ماتریس m×m و A_{ij}، زیر ماتریس گرفته شده از A باشد که ردیف i و ستون j آن خط خورده باشند $(1 \langle i , j \le m)$، بنابر این اسکالر $M_{ij} = |A_{ij}|$، کهاد A با مرتبه (m -1) گفته می‌شود. اسکالر $C_{ij} = (-1)^{i+j} M_{ij}$، (i , j) امین هم عامل A گفته می‌شود.

قضیه بعدی به ما می‌گوید که دترمینان ماتریس را می‌توان با بسط هم عاملهای هر ردیف یا ستون محاسبه کرد.

قضیه (۱): بسط هم عامل‌ها – اگر A یک ماتریس m×m و C_{ij}، (i ,j) امین هم عامل باشد، داریم:

$$|A| = \sum_{i=1}^{m} = a_{ij}C_{ij} \qquad , \quad j=1,\dots m$$

$$|A| = \sum_{j=1}^{m} = a_{ij}C_{ij} \qquad , \quad i=1,\dots m$$

نکته: بسط هم عامل‌ها مخصوص هیچ ردیف یا ستون خاصی نیست. قضیه می‌گوید می‌توانیم هر ردیف یا ستونی را استفاده کنیم که دترمینان A را بر حسب بسط هم عامل‌ها نشان دهد.
خاصیت زیر برای هم عامل‌ها، قضیه بسط خارجی هم عامل‌ها گفته می‌شود.

قضیه (۲): بسط خارجی هم عامل ها – اگر A یک ماتریس m×m و C_{ij} ‚(i , j) امین هم عامل باشد، داریم:

اگر $\sum_{i=1}^{m} a_{ij}C_{ik}=0 \qquad k \neq j$

اگر $\sum_{j=1}^{m} a_{ij}C_{kj}=0 \qquad k \neq i$

قضیه (۳): خواص – اگر A و B ماتریس های m×m باشند، داریم:

الف) حتی اگر A غیر متقارن باشد، $|A| = |A'|$

ب) $|AB| = |A|\,|B|$

پ) به طور کلی، $|A + B| \neq |A| + |B|$

ت) اگر، $A = \mathrm{diag}(a_{11}, \ldots, a_{mm}) = \begin{bmatrix} a_{11} & o & o & o \\ o & a_{22} & & o \\ o & & \ddots & o \\ o & o & o & a_{mm} \end{bmatrix}$

خواهیم داشت: $|A| = a_{11} \ldots a_{mm}$

و به طور خاص، برای همه mها، $|I_m| = 1$

ث) تعویض هر کدام از دو ردیف یا ستونA، علامت دترمینان A را تغییر می‌دهد.

ج) اگر ماتریس A، دو ردیف یا دو ستون یکسان داشته باشد، $|A| = 0$

چ) اگر ماتریس A، یک ردیف یا ستون صفر داشته باشد، $|A| = 0$

ح) اگر λ یک اسکالر باشد و $B=[a_1,\ldots,\lambda\, a_j,\ldots,a_m]$، پس $|B| = \lambda\,|A|$ و اگر $B = \lambda A$ باشد، $|B| = \lambda^n\,|A|$ خواهد بود.

خ) اضافه کردن یک مضرب از ردیف (ستون) s ام از ماتریس A به ردیف (ستون) r ام، دترمینان A را تغییر نمی‌دهد.

د) اگر ردیف‌ها (ستون‌ها)ی ماتریس A، وابستگی خطی داشته باشند، $|A| = 0$.

ر) اگر A یک ماتریس m×m باشد، آنگاه $r(A) = m$،اگر و فقط اگر $|A| \neq 0$ باشد. همچنین A غیر منفرد است، اگر و فقط اگر $|A| \neq 0$ باشد.

توجه: اگر ماتریس A، m×n و دارای r (A) = m با $r \le \min(m,n)$ باشد، بنابر این باید حداقل یک زیر ماتریس مربع از مرتبه r وجود داشته باشد که غیر منفرد باشد و در نتیجه یک دترمینان غیر صفر داشته باشد. تمام ماتریس‌های مربع از مرتبه ... , r + 2 , r + 1 دارای دترمینان صفر می‌باشند. قضیه بعدی را می‌توان استفاده کرد تا یک تعریف پیشنهادی از رتبه ماتریس به دست بیاوریم.

قضیه (۴): رتبه ماتریس A (m×n) برابر است با مرتبه بزرگترین زیر ماتریس که دترمینان غیر صفر داشته باشد.

۵-۱- معکوس ماتریس

تعریف (۱): معکوس – اگر A و B ماتریس‌هایی از مرتبه m باشند، به طوری که :

$$AB = BA = I_m$$

پس B، معکوس A گفته می‌شود. معکوس A را به صورت مقابل می‌نویسیم:

$$A^{-1} = [a^{ij}]$$

قضیه (۱): اگر A و B ماتریس هایی از مرتبه m باشند، داریم:

الف) اگر r(A) = m باشد، A قابل معکوس کردن است.

ب) اگر A غیرمنفرد باشد، A قابل معکوس کردن است.

پ) اگر $|A| \neq 0$ باشد، A قابل معکوس کردن است.

نتایج بعدی می‌توانند برای محاسبه معکوس ماتریس استفاده شوند.

تعریف (۲): ماتریس الحاقی– ماتریس الحاقی A(m×m) به صورت زیر تعریف می‌شود:

$$adj(A) = [C_{ij}]'$$

که در اینجا C_{ij}، (i , j) امین هم عامل است.

قضیه (۲): اگر A یک ماتریس مربع غیر منفرد باشد، داریم:

$$A^{-1} = \frac{\text{adj (A)}}{|A|}$$

قضیه (۳)- خواص- اگر A و B ماتریس‌هایی مربع و غیرمنفرد باشند، داریم:

الف) $(A^{-1})^{-1} = A$

ب) $(AB)^{-1} = (B)^{-1}(A)^{-1}$

پ) $(A')^{-1} = (A^{-1})'$

ت) $|A^{-1}| = \frac{1}{|A|}$

ث) $(A+B)^{-1} \neq A^{-1}+B^{-1}$

به طور کلی ،

حالا نتایجی را در مورد راه حل دستگاه معادلات غیر همگن ارائه می‌دهیم.

قضیه(۴): قانون کرامر- اگر A یک ماتریس $m \times m$ با $r(A) = m$ و $b \neq 0$ یک بردار $m \times 1$ باشد و A_j نمایانگر ماتریس به دست آمده از جا به جایی ستون j با بردار b باشد، بنابر این راه حل منحصر به فرد برای متغیر j ام در دستگاه معادلات ناهمگن $AX = b$ می‌تواند به صورت مقابل محاسبه شود.

$$X_j = \frac{|A_j|}{|A|}$$

۶-۱- ماتریس‌های افراز شده

ماتریس‌های $m \times n$ و $r \times s$ افراز شده زیر را در نظر بگیرید:

$$A=\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \quad , \quad B=\begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

به طوری که A_{11} و B_{11} ماتریس‌های $m_1 \times n_1$ و $r_1 \times s_1$ و A_{12} و B_{21} ماتریس‌های $m_1 \times (n-n_1)$ و $r \times (s-s_1)$ باشند و A_{21} و B_{21} ماتریس‌های $(m-m_1) \times n_1$ و $(r-r_1) \times s_1$ و A_{22} و B_{22} ماتریس‌های $(m-m_1) \times (n-n_1)$, $(r-r_1) \times (s-s_1)$ باشند، سپس نتایج زیر را در مورد عملیات ماتریس‌های افراز شده به دست می‌آوریم:

قضیه (۱): قوانین

الف) جمع $(n_1 = s_1 \quad m_1 = r_1 \ , \ n = s \ , \ m = r)$

$$A+B=\begin{bmatrix} A_{11}+B_{11} & A_{12}+B_{12} \\ a_{21}+B_{21} & A_{22}+b_{22} \end{bmatrix}$$

ب) ضرب $(n_1 = r_1 \ , \ n = r)$

$$A.B=\begin{bmatrix} A_{11}B_{11}+A_{12}B_{21} & A_{11}B_{12}+A_{12}B_{22} \\ A_{21}B_{11}+A_{22}B_{21} & A_{21}B_{12}+A_{22}B_{22} \end{bmatrix}$$

ج) دترمینان $(m = n \ , \ m_1 = n_1)$

به طوری که A_{22} یک ماتریس مربع با دترمینان غیرصفر باشد، $|A| = |A_{22}| \left|A_{11} - A_{12}A_{22}^{-1}A_{21}\right|$ به طوریکه A_{11} یک ماتریس مربع با دترمینان غیر صفر باشد، $|A| = |A_{11}| \left|A_{22} - A_{21}A_{11}^{-1}A_{12}\right|$ به طور اخص، اگر چنانچه A_{12} یا A_{21} ماتریس‌های پوچ باشند، بنابر این:

$$|A| = |A_{11}| \ |A_{22}|$$

د) معکوس $(m_1 = n_1 \ , \ m = n)$

اگر A_{11} و A_{22} ماتریس های مربع غیرمنفرد باشند، داریم:

$$A^{-1}=B=\begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

$$B_{11}=\left[A_{11}-A_{12}A_{22}^{-1}A_{21}\right]^{-1}$$

$$B_{12}=-A_{11}^{-1}A_{12}\left[A_{22}-A_{21}A_{11}^{-1}A_{12}\right],$$

$$B_{21}=-A_{22}^{-1}\left[A_{11}-A_{12}A_{22}^{-1}A_{21}\right]^{-1}$$

$$B_{22}=\left[A_{22}-A_{21}A_{11}^{-1}A_{12}\right]^{-1}$$

توجه: شرح بالا در مورد جمع و ضرب ماتریس‌های افراز شده می‌تواند به ماتریس‌های افراز شده دیگر تعمیم داده شود.

۷-۱- اثر ماتریس

تعریف (۱): اثر – اگر A ماتریس مربع با مرتبه m باشد، آنگاه اثر A که به صورت tr (A) نشان داده می‌شود، عبارت است از مجموع درایه‌های قطر اصلی A، یعنی:

$$tr\,(A)=\sum_{i=1}^{m} a_{ij}$$

قضیه (۱): قوانین

الف) چنانچه A و B ماتریس‌های مربع باشند، سپس،

$$tr\,(A+B)=tr\,(A)+tr\,(B)$$

ب) چنانچه A و B به ترتیب ماتریس‌هایی از بعد m×n و n×m باشند، سپس:

$$tr(AB)=tr(BA)$$

مثال :

$$trx'Ax=tr(xx'A)=tr\,(Axx')$$

۸-۱- مقادیر ویژه (ریشه های ویژه) و بردارهای ویژه

تعریف(۱): ریشه ها و بردارهای ویژه- چنانچه A یک ماتریس مربع m × m و λ ریشه ویژه A اسکالر و x یک بردار غیر هیچ m×1 باشد، اگر Ax = λx باشد، پس λ ریشه ویژه A (eigen value) گفته می‌شود و x، بردار ویژه آن (eigen vector) خواهد بود.

توجه: البته، شرط Ax = λx برابر است با شرط $(A-\lambda I)x=0$. احتیاج داریم که $x \neq 0$ و پس، ستون‌های A - λI باید وابستگی خطی داشته باشند. پس، می‌توانیم همه λهایی را که در معادله $(A-\lambda I)=0$ صدق می‌کنند برای x های مخالف صفر، از طریق حل معادله زیر برای λها، پیدا کنیم:

$$|A-\lambda I|=0$$

به این معادله، معادله ویژه گفته می‌شود. به طور کلی، ملاحظه می‌شود که معادله ویژه، یک چند جمله‌ای بر حسب λ از درجه m می‌باشد، که m بعد ماتریس A می‌باشد.

$$\lambda^m+P_{m-1}\lambda^{m-1}+\cdots+P_0=0 \qquad , \qquad P_0=|-A|$$

از قضایای اساسی جبر غیرخطی می‌دانیم که یک چند جمله‌ای از درجه m داری m ریشه $\lambda_1\ldots\lambda_m$ می‌باشد، بنابر این برای محاسبه ریشه‌های ویژه ماتریس A باید ریشه‌های معادله ویژه را برای λ_iها حل کنیم. چون ستون‌های $A-\lambda_i I$ به طور خطی به هم وابسته‌اند، حداقل یک بردار غیر صفر مانند x_i وجود دارد، به طوری که:

$$(A-\lambda_i I)x_{.i}=0$$

بنابر این برای به دست آوردن بردار ویژه متناظر با ریشه ویژه، می‌باید دستگاه m معادله خطی را حل کنیم.

قضیه (۱): قوانین- چنانچه A ماتریس $m\times n$ با ریشه های ویژه $\lambda_1\ldots\lambda_m$ باشد، سپس داریم:

الف) $|A|=\prod_{i=1}^{m}\lambda_i$

ب) $tr(A)=\sum_{i=1}^{m}\lambda_i$

پ) ریشه‌های ویژه A' نیز عبارتند از $\lambda_1\ldots\lambda_m$، یعنی A و A' دارای ریشه‌های ویژه یکسان‌اند.

ت) اگر ماتریس A غیر منفرد باشد، کلیه ریشه‌های ویژه آن غیرصفر می‌باشند و ریشه‌های ویژه A^{-1}، $\mu_1=\frac{1}{\lambda_1},\cdots,\mu_m=\frac{1}{\lambda_m}$ خواهند بود.

ث) ریشه‌های ویژه A^2، $\lambda_1^2,\ldots\lambda_m^2$ خواهند بود.

ج) چنانچه B ماتریس $m\times n$ باشد، آنگاه ریشه‌های ویژه AB دقیقا برابر با ریشه های ویژه $B\times A$ خواهد بود.

تعریف (۲): ریشه های تکراری مجزا و تکراری – چنانچه A ماتریس $m\times m$ باشد، فرض کنید که:

$$|A-\lambda I|=\prod_{i=1}^{s}(\lambda-\lambda_i)^{m_i}$$

به طوری که برای $i\neq j$, $\sum_{i=1}^{s}m_i=m$, $\lambda_i\neq\lambda_j$ باشد.

سپس می‌گوییم ماتریس A دارای ریشه‌های مجزای $\lambda_1,\ldots\lambda s$ و می‌باشد و ریشه‌های λ_i , m_i بار تکرار شده است.

قضیه (۲): چنانچه A ماتریس $m \times m$ و $\lambda_1, \ldots \lambda_s$ ریشه‌های ویژه مجزای A و $x_1, \ldots, x_s$ به ترتیب بردارهای ویژه مربوط به $\lambda_1, \ldots, \lambda_s$ باشند، آنگاه بردارهای $X_1, \ldots, X_s$ بطور خطی مستقل‌اند.

قضیه (۳): قطری کردن- چنانچه A ماتریس $m \times m$ باشد و فرض کنیم همه ریشه‌های ویژه‌اش ($\lambda_1, \ldots, \lambda_m$) مجزا باشند و چنانچه $x_1, \cdots, x_m$ بردار ویژه مربوط به ریشه‌های ویژه $\lambda_1, \ldots, \lambda_m$ و $X = [x_1, \cdots, x_m]$ و Λ ماتریس قطری با درایه‌های $\lambda_1, \ldots, \lambda_m$ بر روی قطر اصلی باشند، آنگاه ماتریس A را می‌توان به صورت زیر قطری کرد.

$$X'AX = \Lambda$$

به عبارت دیگر ماتریس A شبیه ماتریس قطری است.

۹-۱- ماتریس‌های متعامد

تعریف (۱): ماتریس‌های متعامد

الف- اگر a و b بردارهای $m \times 1$ باشند، آنگاه a و b متعامد گفته می‌شوند، اگر و فقط اگر $a'b = 0$. بعلاوه، اگر $a'a = 1$ و $b'b = 1$ باشد، متعامد نرمال گفته می‌شوند.

ب- اگر Q ماتریس $m \times m$ باشد، آنگاه Q متعامد خوانده می‌شود، اگر و فقط اگر $Q'Q = I$ باشد، یعنی اگر و فقط اگر ستونهای Q متعامد باشند.

قضیه (۱): خواص- چنانچه Q ماتریس $m \times m$ و $\lambda_1, \ldots, \lambda_m$ نشان دهنده مقادیر ویژه آن باشد، سپس:

الف- Q غیر منفرد است

ب- $Q^{-1} = Q'$ (و همچنین $Q'Q = I$)

پ- $|Q| = \pm 1$

ت- $\lambda_i = \pm 1, (i = 1, \ldots n)$

توجه: چنانچه $g_1^*, g_2^*, \cdots, g_m^*$ مجموعه بردارهای مستقل از هم باشند، آنگاه می‌توان با استفاده از روش Gram – Schmidt، آنها را به مجموعه متعامد نرمال تبدیل کرد.

۱۰-۱- ماتریس‌های متقارن

قضیه (۱): چنانچه S ماتریسی حقیقی از مرتبه m باشد که درایه‌های آن حقیقی باشند، آنگاه ریشه‌های ویژه آن نیز حقیقی می‌باشند.

قضیه (۲): چنانچه S ماتریسی حقیقی و متقارن از مرتبه m بوده و ریشه‌های مجزای آن $\lambda_i , i=1,\cdots S\leq m$ باشند، هم‌چنین حاصل ضرب m_i , λ_i باشد، بطوریکه $\sum_{i=1}^{s} m_i = m$ باشد، سپس متناظر با هر m_i , λ_i بردار ویژه متعامد نرمال (خطی مستقل) وجود دارد.

قضیه (۳): چنانچه S یک ماتریس حقیقی از مرتبه m باشد، در نتیجه بردارهای ویژه S می‌توانند به عنوان یک مجموعه متعامد نرمال انتخاب شوند.

نتیجه (۱)- چنانچه S یک ماتریس حقیقی متقارن از مرتبه m باشد و $\lambda_1, \ldots, \lambda_m$ (احتمالا $\lambda_i = \lambda_j$ خواهد بود) نمایانگر ریشه های ویژه و $q_{.1},\cdots,q_{.m}$، مجموعه‌ای متناظر با بردارهای ویزه متعامد نرمال باشد و $Q=[q_{.1},\cdots,q_{.m}]$، ماتریس‌های متعامد تشکیل یافته از آن بردارهای ویژه باشد و $\Lambda = diag\,(\lambda_1,\cdots,\lambda_m)$ داریم:

$$Q'SQ=\Lambda \quad , \quad S=Q\Lambda Q'$$

یعنی، می توانیم ماتریس S را از طریق پیش ضرب کردن Q' و Q قطری کنیم، به عبارتی دیگر S متعامدا" به یک ماتریس قطری شبیه است.

توجه: برای این که ببینیم نتیجه (۱) چگونه از قضیه فوق استنباط می‌شود، مشاهده می‌کنیم که طبق تعریف $q_{.1}, \lambda_i$، داریم $Sq_{.i} - q_{.i}\lambda_i = 0$ به طوری که $i=1,\cdots,m$ و یا به طور بسته ماتریسی داریم: $SQ-Q\Lambda=o$ آنگاه حاصل با پیش ضرب نمودن Q' و با توجه به این که $Q'Q=I$ می‌باشد، به دست می‌آید. چون Q غیرمنفرد است، همچنین $r\,(Q'SQ)=r\,(\Lambda)$ می باشد و نتیجه زیر به دست می‌آید.

قضیه (۴): چنانچه S ماتریسی حقیقی و متقارن با مرتبه m باشد، آنگاه:

$r\,(S)=r\,(\Lambda)=$ تعداد ریشه های ویژه غیر صفر

۱-۱۱- ماتریس هم قوه

تعریف (۱): ماتریس هم قوه – اگر ماتریس A با مرتبه m × m را در نظر بگیریم، آنگاه A هم قوه خواهد بود، اگر و فقط اگر AA=A باشد.

قضیه (۱): اگر A ماتریسی m × m و هم قوه باشد، آنگاه:

الف- ریشه ویژه، A یا صفر است یا یک.

ب- tr(A) = r (A)

۱-۱۲- ماتریس های شبه معین و معین

تعریف (۱): شکل درجه دوم- چنانچه $A=[a_{ij}]$، ماتریسی m × m و x برداری m × 1 باشد، آنگاه تابع $f : R^m \rightarrow R$ به صورت $f\,(X)=X'Ax=\sum_{i=1}^{m}\sum_{j=1}^{m} a_{ij} X_i X_j$ شکل درجه دوم A خوانده می‌شود.

چون $X'AX=(\frac{1}{2})X'(A+A')X$ فرض می‌کنیم که از این پس ماتریس A متقارن است، این فرض هیچ گونه مشکلی ایجاد نخواهد کرد.

تعریف (۲): اجازه دهید A ماتریسی m×m و متقارن باشد، آنگاه:

الف – A مثبت شبه معین است، اگر برای همه x ها، $X'AX \geq 0$ باشد.

ب- A مثبت معین است، اگر برای همه $x\neq 0$ ها، $x'Ax \rangle 0$ باشد.

پ- A منفی شبه معین است، اگر A– مثبت شبه معین باشد.

ت- A منفی معین است، اگر A– مثبت معین باشد.

قضیه (۱): اجازه دهید $A=[a_{ij}]$ ماتریس m ×m متقارنی باشد با ریشه‌های ویژه $\lambda_1,\lambda_2,\dots,\lambda_m$

الف- اگر A مثبت معین باشد، آنگاه A مثبت شبه معین نیز هست.

ب- اگر A مثبت معین باشد آنگاه به ازای $a_{ij} \rangle 0,, i=1,\dots m$ خواهد بود.

اگر A مثبت شبه معین باشد، آنگاه به ازای $a_{ij} \geq 0,, i=1,\dots m$ خواهد بود.

پ-A مثبت معین است، اگر و فقط اگر:

$$a_{11}\rangle 0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \rangle 0 \ , \ \cdots , \det \begin{vmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mm} \end{vmatrix} = |A| \rangle 0.$$

ت- A مثبت معین است، اگر و فقط اگر به ازای $\lambda_i \rangle 0,, i=1,\dots m$ باشد.

A مثبت شبه معین است، اگر و فقط اگر به ازای $\lambda_i \geq 0,, i=1,\dots m$ باشد.

ث- اگر A مثبت شبه معین باشد، اما مثبت معین نباشد، آنگاه r(A) < m.

ج- اگر A مثبت شبه معین باشد، آنگاه ماتریسی m×m به نام S با رتبه m وجود دارد، به طوری که $A=S'S$.

اگر A مثبت شبه معین باشد، آنگاه ماتریس S (m×m) دارای رتبه کوچکتر از m است، به طوری که $A=S'S$.

چ-چنانچه $S_{m\times m}$ باشد و $A_{m.m}=S'S$، آنگاه اگر r(s) = m×m باشد، A مثبت معین است و اگر A, r(s)< m مثبت شبه معین است (نه مثبت معین)

ح- چنانچه A و B مثبت معین باشند و اگر B-A مثبت معین باشد، آنگاه A-1 – B-1 مثبت معین است.

تمریـــــــــن

سئوال اول – ماتریس مقابل را در نظر بگیرید:

$$x=\begin{bmatrix} 1 & 1 & 1 & 3 \\ 1 & 2 & 0 & 5 \\ 1 & 1 & 1 & 2 \end{bmatrix}$$

الف- ابعاد ماتریس x چیست؟

ب- $x'x$ را محاسبه کنید. راجع به ویژگی‌های حاصل ضرب چه می‌توان گفت؟

(یعنی آیا حاصل ضرب، مربع، متقارن، غیرمنفرد و ... است).

پ- حال ماتریس $A=[I_n-(x'x)]$ را در نظر بگیرید.

۱- مقدار n در معادله چیست؟

۲- A را محاسبه کنید.

۳- فرض کنید A را به صورت $A_{11},A_{12},A_{21},A_{22}$ افراز کرده‌ایم، به طوری که:

$$A=\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{21} \end{bmatrix}$$

و می‌دانیم A_{11} ,3×1 و 1×1 ,A_{21}است. ابعاد A_{21} چه خواهد شد؟ بر حسب مقادیر به دست آمده در قسمت ۲، زیر ماتریس‌های $A_{11},A_{12},A_{21},A_{22}$ را بنویسید.

۴- آیا A هم قوه است؟

۵- نشان دهید که $X'X$ غیر منفرد است و B را به صورت $I-X\,(X'X)^{-1}X$ تعریف کنید. آیا B هم قوه است؟

سئوال دوم- بردار $X=\begin{bmatrix}1\\1\\1\end{bmatrix}$ را در نظر بگیرید و $X'X$ و XX' را محاسبه کنید.

سئوال سوم- آیا به طور کلی $(AB)^k=B^K A^K$؟ اگر نه، چه شرایطی باید بر روی A و B قرار دهیم تا رابطه فوق برقرار باشد؟

سئوال چهارم-

الف- آیا بردارهای $a_1=\begin{bmatrix}1\\2\end{bmatrix}$, $a_2=\begin{bmatrix}7\\a\end{bmatrix}$, $a_3=\begin{bmatrix}1\\7\end{bmatrix}$ به طور خطی مستقل از یکدیگرند؟

ب- در ماتریس $A=[a_1\ a_2\ a_3]$، رتبه ستونی A چیست؟

پ- رتبه سطری آن چیست؟

ت- آیا A دارای رتبه کامل است؟

ث- حال، ماتریس زیر را در نظر بگیرید:

$$B=\begin{bmatrix}1 & 8\\ 9 & 3\end{bmatrix}$$

رتبه BA را به دست آورید.

سئوال پنجم- ماتریس بالا مثلثی زیر را در نظر بگیرید:

$$A=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ o & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ o & o & \cdots & a_{nn} \end{bmatrix}$$

نشان دهید که $|A| = a11a22\cdots av\mu$

یعنی $|A|$ برابر حاصل ضرب درایه‌های روی قطر است.

سئوال ششم-

الف- دستگاه معادلات زیر را به صورت نماد ماتریسی Ax = b بنویسید:

$$8x_1 - x_2 = 15$$
$$x_2 - 5x_3 = 1$$
$$2x_1 + 3x_3 = 4$$

ب- معکوس ماتریس ضرایب A را به دست آورید و جواب x^* را با استفاده از فرمول $x^* = A^{-1}b$ به دست آورید.

پ- فرض کنید b = 0 است. حالا x^* را به دست آورید، آیا x^* جواب منحصر به فرد است؟ چگونه می‌توان تشخیص داد؟

سئوال هفتم – تابع عرضه و تقاضا در دو بازار رقابتی را در نظر بگیرید:

برای بازار اول $Q_d^1 = 10 - p^1 + p^2 \quad , \quad Q_s^1 = 1 + 2q^1$

برای بازار دوم $Q_d^2 = 20 + 2p^1 - 4p^2 \quad , \quad Q_s^2 = p^2$

الف- چنانچه Q1 , Q2 نشان دهنده مقادیر تعادلی در دو بازار باشد، مدل دو بازار فوق را در قالب Ax=d بنویسید، به طوری که $x = [Q^1, Q^2, P^1, P^2]$

ب- با استفاده از روش کرامر قیمتهای تعادلی را به دست آورید.

پ- مقادیر تعادلی را با هر روشی که می‌دانید به دست آورید.

سئوال هشتم – ماتریس $A = \begin{bmatrix} 2 & 2 \\ 2 & - \end{bmatrix}$ را در نظر بگیرید.

الف- معادله ویژه ماتریس فوق را بنویسید. مقادیر ویژه مربوط به آن را به دست آورید.

ماتریس قطری L را ایجاد نمایید، به طوری که مقادیر ویژه درایه‌های قطر اصلی آن باشند.

ب- بردارهای ویژه متناظر با مقادیر ویژه را به دست آورید.

پ- مقادیر ویژه را طوری نرمالیزه کنید که متعامد نرمال شوند.

ت- با استفاده از بردارهای نرمالیزه شده، ماتریس متعامد X را ایجاد نمایید. بررسی کنید که X واقعا متعامد است.

ث- ماتریس Λ را محاسبه کنید، به طوری که $\Lambda = X'AX$ باشد . این ماتریس را با L مقایسه نمایید.

ج- با استفاده از اطلاعاتی که راجع به مقادیر ویژه دارید، می‌توانید راجع به معین بودن ماتریس A و در نتیجه درباره فرم درجه دوم $q(x) = X'AX$ اظهار نظر کنید.

سئوال نهم – فرم درجه دوم زیر را در نظر بگیرید.

$$q(x_1,,x_2,x_3) = 4(x_1^2 + x_2^2 + x_3^2 + x_1x_2 + x_1x_3 + x_2x_3)$$

الف- فرم بالا را به صورت ماتریس زیر بنویسید.

$$q(x) = x'Ax$$

ب- این فرم را از نظر معین یا نامعین بودن با استفاده از زیر دترمینان‌های آن، بررسی کنید.

سئوال دهم- فرض کنید در سئوال چهارم به جای ماتریس A، از ما خواسته شده است که آزمون مشابهی برای ماتریس W انجام دهیم. آیا می‌توانید قضیه مشابهی را که در سئوال چهارم به کار می‌بریم، استفاده کنیم؟ اگر جواب منفی است، می‌توانید با تبدیل W، قضیه مورد نظر را برای آن به کار ببرید.

ماتریس مبدل چگونه خواهد بود؟

$$W = \begin{bmatrix} 0 & 5 & 1 \\ 6 & 3 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

منابع:

1- Amemiya, T.”Advanced Econometrics”, Cambrigde: Harvard University Press, 1985.
2- Baltaghi Badi H. “Econometrics” 2nd ed. Springer, 1999.
3- Breusch, T., and A.Pagan.”Asimple Test for Hetroscedasticity and Random Coefficient Variation.” Econometrica, 47, 1977, pp.1287-94.
4- Cochran, D., and G. Orcutt. ”Application of Least Squares Regression to Relationships Containing Autocorrelation Error Terms.”Journal of the American Statistical Association, 44, 1949, pp.32-61.
5- Dickey, D., and W. Fuller. “Distribution of the Estimators for Autoregressive Time Series with a Unit Root.” Journal of the American Statistical Association, 74, 1979, pp.427-431.
6- Dickey, D., and W. Fuller. “Likelihood Ratio Testsfor Autoregressive time Series with a Unit Root.” Econometrica, 49, 1981, pp.1057-1072.
7- Diebold, F., and M. Nerlove. “Unit Roots in Econometric Time Series: A Selective Survey.” In T. Bewley, ed., Advances in Econometrics, Vol. 8.New York: JAI Press, 1990.
8- Fomby Thomas B, R.Carter Hill, Stanley R. Johnson “Advanced Econometric Methods” Springer-Verlag 1984.
9- Greene William H. “Econometrics Analysis” 3rd, ed. 2000.
10- Maddala G.S, “Introduction to Econometrics” Macmillan Publishing Company, 3rd ed.1999.
11- Theil, H. “Principles of Econometrics”Wiley, 1971.
12- Durbin, J., and G. Watson. “Testing for Serial Correlation in Least Squares Regression-I.” Biometrika, 37, 1950, pp.409-428.
13- Durbin, J., and G. Watson. “Testing for Serial Correlation in Least Squares Regression-II.” Biometrika, 38, 1951, pp.159-178.
14- Durbin, J., and G. Watson. “Testing for Serial Correlation in Least Squares Regression-II.” Biometrika, 58, 1971, pp.1-42.
15- Engle, R., and C Granger. “Co-integration and Error Correction: Representation, Estimation and Testing.” Econometrica, 35, 1987, pp.251-276.
16- Frisch, R. “Editorial.” Economerrica, 1, 1993, pp. 1-4.
17- Glesjer, H. “A New Test for Heteroscedasticity.” Journal of the American Statistical Association, 64, 1969, pp. 316-323.
18- Godfrey, L. “Testing Against General Autoregressive and Moving Average Error Models When the Regressors Include Lagged Dependent Variables.” Economiertica, 46. 1978, pp. 1293-1302.
19- Goldberger, A. “Best Linear unbiased Prediction in the Generalized Regression Model.” Journal of the American Statistical Association, 57, 1962, pp. 369-375.
20- Goldfeld, S., and R. Quandt. “Some Tests for Homoscedasticity.” Journal of the American Statistical Association, 60, 1965, pp. 539-547.
21- Gujarati, D. Basic Econometrics, 3rd Ed. New York, McGraw-Hill, 1995.
22- Hamilton, J. Time Series Analysis. Princeton: Princeton University Press, 1994.
23- Harvey, A. “Estimating Regression Models with Multiplicative Heteroscedasticity.” Econometrica, 44, 1978, pp. 465.
24- Johnston, J. Econometric Methods. New York: McGraw-Hill, 1984.

25- Johnston, J. and J. DiNardo, Econometric Methods, 4^{th} Ed. New York: McGraw-Hill, 1997.
26- Judge, G., C.Hill, W. Griffiths, and T. Lee. The Theory and Practice of Econometrics. New York: John Wiley and Sons, 1985.
27- Kelejian, H. “Two-Stage Least Squares and Econometric Systems Linear in Parameters but Nonlinear in the Endogenous Variables.” Journal of the American Statistical Association, 66, 1971, pp. 373-374.
28- Keynes, J. The General Theory of Employment, Interest, and Money. New York: Harcourt, brace, and Jovanovich, 1936.
29- Lucas, R. “Econometric Policy Evaluation: A Critique.” In K. Brunner and A. Meltzer, eds., The Philips Curve and the labor Market. Amsterdam: North Holland, 1976.
30- Nelson, C., and C. Plosser. : Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications.” Journal of Monetary Economics, 10, 1982, pp.139-162.
31- Phillips, P., and P. Perron. “Testing for a Unit Root in Time Series Regression.” Biometrika, 75, 1988, pp. 335-346.
32- Zellner, A. “Estimators for Seemingly Unrelated Regression Equations: Some Exact Finite Sample Results.” Journal of the American Statistical Association, 58, 1963, pp. 977-992.