

حاشا
الرحمن الرحيم



دانشکده علوم ریاضی

رشته آمار، گرایش آمار ریاضی

پایان نامه کارشناسی ارشد

مدل معادلات ساختاری ریج

نگارنده: علی سلاجقه

استاد راهنما

دکتر محمد آرشی

شهریور ۱۳۹۶



فرم شماره ۶: صورتجلسه دفاع از پایان نامه تحصیلی دوره کارشناسی ارشد

با تأییدات خداوند متعال و با استعانت از حضرت ولی عصر (عج) ارزیابی جلسه دفاع از پایان نامه کارشناسی ارشد **محمدعلی...** به شماره دانشجویی **۱۳۹۴۱۰۲۷** رشته ریاضی کاربردی گرایش آنالیز عددی تحت عنوان **همگرایی بدون مش...** برای معادلات سهموی نام غیرمطلوبه که در تاریخ ۱۳۹۴/۱۰/۲۷ با حضور هیأت محترم داوران در دانشگاه صنعتی شاهرود برگزار گردید به شرح ذیل اعلام می گردد:

قبول (با درجه: **بسیار امتیاز (۱۸)**) دفاع مجدد مردود

۲- بسیار خوب (۱۸-۱۸/۹۹)

۱- عالی (۲۰-۱۹)

۴- قابل قبول (۱۴-۱۵/۹۹)

۳- خوب (۱۶-۱۷/۹۹)

۵- نمره کمتر از ۱۴ غیر قابل قبول

امضاء	مرتبه علمی	نام و نام خانوادگی	عضو هیأت داوران
	استادیار	دکتر علی مس فروش	۱- استاد راهنمای اول
			۲- استاد راهنمای دوم
	استادیار	دکتر مهدی قوتمند	۳- استاد مشاور
	استاد	دکتر ابراهیم هاشمی	۴- نماینده شورای تحصیلات تکمیلی
	استادیار	دکتر حجت احسنی طهرانی	۵- استاد ممتحن اول
	دانشیار	دکتر علیرضا ناظمی	۶- استاد ممتحن دوم

رئیس دانشکده:

تقدیم بہ روح پاک محبوبہ جان خواہرم
کہ وجودش شادی بخش و صفایش مایہ آرامش من بود.

سپاس‌گزاری...

با سپاس از سه وجود مقدس:
آنان که ناتوان شدند تا ما به توانایی برسیم...
موهایشان سپید شد تا ما روسفید شویم...
و عاشقانه سوختند تا گرمابخش وجود ما و روشن‌گر راهمان باشند...
پدرانمان
مادرانمان
استادانمان

برخود واجب می‌دانم مراتب سپاس و قدردانی عمیق قلبی خود را به خدمت استاد فرزانه و گران‌قدم جناب آقای دکتر محمد آرشی که در طول دوران تحصیل و ارائه پایان‌نامه از چشمه جوشان دانش و اخلاق والایشان بهره‌مند گشته‌ام، ابراز نمایم. هم‌چنین از اساتید بزرگوار و دلسوز دکتر محمدرضا ربیعی و دکتر سیدرضا حسینی‌نیا که زحمت داوری این رساله را متقبل شدند؛ کمال تشکر و قدردانی را دارم. هم‌چنین از اساتید محترم خانم دکتر اقبال، دکتر شاهسونی، دکتر باغیشنی و دکتر بلبلیان که در طول دوران تحصیل مقطع کارشناسی ارشد که دانش و اندوخته علمی خودشان را در اختیارم قرار دادند تشکر و قدردانی می‌کنم.

هم‌چنین از زحمات بی‌شائبه پدر و مادر عزیز و همسر مهربانم تشکر می‌کنم که اگر کمک این عزیزان نبود هرگز به این مهم دست نمی‌یافتیم. و در آخر از تمام کسانی که در این راه به بنده کمک کرده‌اند بالاخص خواهران دلسوز و برادر عزیزم و آقای دکتر مهدی اکبرزاده بسیار سپاسگزارم.

علی سلاجقه

شهریور ۱۳۹۶

تعهد نامه

اینجانب علی سلاجقه دانشجوی کارشناسی ارشد رشته آمار علوم ریاضی دانشگاه شاهرود، نویسنده پایان نامه با عنوان **مدل معادلات ساختاری ریج**، تحت راهنمایی محمد آرشی متعهد می شوم:

- تحقیقات در این پایان نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است.
- در استفاده از نتایج پژوهش های دیگر پژوهش گران، به مرجع مورد استفاده استناد شده است.
- مطالب این پایان نامه، تا کنون توسط خود، یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ جا ارایه نشده است.
- حقوق معنوی این اثر، به دانشگاه صنعتی شاهرود تعلق دارد، و مقالات مستخرج با نام “ دانشگاه صنعتی شاهرود “ یا “ Shahrood University of Technology “ به چاپ خواهد رسید.
- حقوق معنوی تمام افرادی که در به دست آوردن نتایج اصلی پایان نامه تاثیرگذار بوده اند، در مقالات مستخرج از پایان نامه رعایت می گردد.
- در تمام مراحل انجام این پایان نامه، در مواردی که از موجود زنده (یا بافت های آنها) استفاده شده است، ضوابط و اصول اخلاقی رعایت شده است.
- در تمام مراحل انجام این پایان نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته (یا استفاده شده است)، اصل رازداری و اصول اخلاق انسانی رعایت شده است.

علی سلاجقه

شهریور ۱۳۹۶

مالکیت نتایج و حق نشر

- تمام حقوق معنوی این اثر و محصولات آن (مقالات مستخرج، کتاب، برنامه های رایانه ای، نرم افزارها و تجهیزات ساخته شده) متعلق به دانشگاه صنعتی شاهرود می باشد. این مطلب باید به نحو مقتضی، در تولیدات علمی مربوطه ذکر شود.
- استفاده از اطلاعات و نتایج موجود در این پایان نامه بدون ذکر منبع مجاز نمی باشد.

چکیده

در برخی مطالعات متغیرهایی وجود دارد که نمی‌توان به‌طور مستقیم آن‌ها را اندازه‌گیری یا مشاهده کرد و اغلب به آن‌ها متغیرهای پنهان می‌گویند. مدل‌سازی معادلات ساختاری یک روش مستقیم برای مدل‌سازی متغیرهای پنهان فراهم می‌کند، که ترکیب مدل ساختاری و مدل اندازه‌گیری، مدل‌سازی معادلات ساختاری را تشکیل می‌دهد. زمانی که بین متغیرهای پیشگو همبستگی شدیدی وجود داشته باشد گوییم همخطی چندگانه وجود دارد. با وجود مشکل همخطی چندگانه برآوردگر کمترین توان‌های دوم کاربردی نبوده و اغلب از برآوردگرهای انقباضی اریب استفاده می‌شود. روش‌های متعددی برای به‌دست آوردن برآوردگرهای انقباضی ضرایب رگرسیون مطرح شده است. یکی از این روش‌ها استفاده از برآوردگر ریج است. در اینجا روش ریج را برای حل مدل معادلات ساختاری در داده‌های رسته‌ای و پیوسته مورد بررسی قرار دادیم. در این راستا به عناصر روی قطر اصلی ماتریس همبستگی یک مقدار ثابت اضافه نموده و سپس تابع هدف در مدل معادلات ساختاری را بر پایه‌ی این تغییر، به‌منظور برآورد ضرایب مدل کمینه کردیم. این روش را در قالب یک مثال عددی مورد بررسی قرار دادیم. در پایان نیز به مدل‌سازی معادلات ساختاری منظم در قالب یک مثال و مقایسه روش ریج با چند روش انقباضی دیگر پرداختیم.

کلمات کلیدی: همخطی چندگانه، متغیر پنهان، مدل‌سازی معادلات ساختاری، برآوردگر ریج، برآوردگرهای انقباضی، ماتریس همبستگی

لیست مقالات مستخرج از پایان نامه

۱. سلاجقه ع. و آرشی م. (۱۳۹۶) تأثیر سرمایه اجتماعی و رضایت شغلی بر تعهد سازمانی کارکنان ادارات ورزش و جوانان استان اصفهان و ارائه مدل، نشریه مدیریت منابع انسانی در ورزش، (در حال داوری).
۲. سلاجقه ع. و آرشی م. (۱۳۹۶) مدل سازی معادلات ساختاری ریج، چهل و هشتمین کنفرانس ریاضی ایران- دانشگاه بوعلی سینا همدان.
۳. سلاجقه ع.، آرشی م. و رحمانی م. (۱۳۹۶) مدل معادلات ساختاری جریمه ریج، کنفرانس روش های مدرن در قیمت گذاری های بیمه ای و آمارهای صنعتی- دانشگاه بوعلی سینا همدان.

فهرست مطالب

ق	فهرست تصاویر
ش	فهرست جداول
۱	۱ مدل سازی معادلات ساختاری
۱	۱.۱ مقدمه
۲	۲.۱ مدل های رگرسیون
۳	۳.۱ رگرسیون خطی چندگانه
۳	۱.۳.۱ برآورد ضرایب رگرسیون
۵	۴.۱ تحلیل عاملی
۸	۱.۴.۱ انواع تحلیل عاملی
۱۲	۲.۴.۱ تحلیل عاملی تاییدی
۱۴	۳.۴.۱ برآورد پارامترها به روش ماکزیمم درستنمایی
۱۵	۵.۱ تحلیل مسیر
۱۷	۱.۵.۱ طراحی مدل مسیر
۱۷	۲.۵.۱ انواع مدل های تحلیل مسیر
۱۸	۳.۵.۱ انواع روابط بین متغیرها در نمودار مدل مسیر
۱۸	۴.۵.۱ مفروضات تحلیل مسیر
۱۹	۵.۵.۱ اصول ترسیم نمودار مسیر
۱۹	۶.۱ مدل سازی معادلات ساختاری
۲۰	۱.۶.۱ الگوی ریاضی مدل معادلات ساختاری
۲۱	۲.۶.۱ انواع متغیرها در مدل یابی معادلات ساختاری
۲۷	۳.۶.۱ مراحل مدل سازی معادلات ساختاری
۳۰	۷.۱ شاخص های برازش
۳۱	۱.۷.۱ شاخص های برازش مطلق
۳۴	۲.۷.۱ شاخص های برازش تطبیقی

۳۷	شاخص‌های برازش مقتصد	۳.۷.۱
۴۰	سایر شاخص‌ها	۴.۷.۱
۴۳	مثال‌های کاربردی	۸.۱
۴۳	مثال مدیریت ورزشی	۱.۸.۱
۵۰	مثال روان‌شناسی	۲.۸.۱
۵۹	مدل معادلات ساختاری ریج	۲
۵۹	مقدمه	۱.۲
۵۹	همخطی در رگرسیون چندگانه	۲.۲
۶۲	برآوردگر ریج	۳.۲
۶۳	مدل معادلات ساختاری ریج	۴.۲
۶۵	خواص برآوردگر ریج در مدل معادلات ساختاری	۵.۲
۶۵	سازگاری	۱.۵.۲
۶۶	نرمال مجانبی	۲.۵.۲
۶۸	آماره‌هایی برای ارزیابی کلی مدل	۳.۵.۲
۷۰	یک مثال واقعی	۶.۲
۷۱	بحث و نتیجه‌گیری	۷.۲
۷۷	مدل معادلات ساختاری جریمه شده	۳
۷۷	مقدمه	۱.۳
۷۸	درست‌نمایی جریمه شده	۲.۳
۷۹	انواع توابع جریمه	۱.۲.۳
۸۱	بسته regsem	۳.۳
۱۰۴	مقایسه	۴.۳
۱۰۷	نتیجه‌گیری	۵.۳
۱۰۹	خلاصه	۶.۳
۱۰۹	پیشنهادات برای آینده تحقیق	۷.۳
۱۱۱	آ خروجی نرم‌افزارها	
	روش اجرای مدل‌سازی معادلات ساختاری در نرم‌افزار Amos برای مثال	۱.آ
۱۱۱	مدیریت ورزشی در فصل ۱	
	روش اجرای مدل‌سازی معادلات ساختاری در نرم‌افزار R برای مثال روان‌شناسی	۲.آ
۱۳۱	در فصل ۱	
	روش اجرای مدل‌سازی معادلات ساختاری ریج در نرم‌افزار SAS برای مثال	۳.آ
۱۳۴	واقعی در فصل ۲	

۱۶۱

مراجع

۱۶۷

واژه‌نامه انگلیسی به فارسی

۱۷۱

واژه‌نامه فارسی به انگلیسی

فهرست تصاویر

۶	شکل‌های مورد استفاده در مدل‌سازی معادلات ساختاری	۱.۱
۷	نمودار رگرسیون چندگانه	۲.۱
۷	نمودار رگرسیون چندمتغیره	۳.۱
۹	نمودار تحلیل عاملی اکتشافی	۴.۱
۹	نمودار تحلیل عاملی تاییدی	۵.۱
۱۵	نمودار مسیر	۶.۱
۱۶	نمودار مسیر با متغیرهای پنهان	۷.۱
۱۹	نمودار مدل اندازه‌گیری تکی	۸.۱
۲۰	نمودار مدل اندازه‌گیری چندگانه	۹.۱
۲۴	نمودار مسیر مدل معادلات ساختاری	۱۰.۱
۲۵	نمودار مدل معادلات ساختاری	۱۱.۱
۲۹	نمودار مدل شناسایی در مدل‌سازی معادلات ساختاری	۱۲.۱
۴۴	مدل مفهومی پژوهش	۱۳.۱
۴۶	مدل‌سازی معادلات ساختاری	۱۴.۱
۸۷	نمودار اثر ریبج داده‌های بخش ۳.۳	۱.۳
۹۱	نمودار اثر لاسو داده‌های بخش ۳.۳	۲.۳
۹۴	نمودار اثر لاسوی تطبیقی داده‌های بخش ۳.۳	۳.۳
۹۷	نمودار اثر الاستیک‌نت داده‌های بخش ۳.۳	۴.۳
۱۰۰	نمودار اثر scad داده‌های بخش ۳.۳	۵.۳
۱۰۳	نمودار اثر mcp داده‌های بخش ۳.۳	۶.۳
۱۰۴	نمودار مدل منحنی رشد پنهان	۷.۳

فهرست جداول

۴۵ ماتریس همبستگی متغیرهای تحقیق	۱.۱
	شاخص‌های برآزش تاثیر متغیر سرمایه اجتماعی و تعهد سازمانی بر رضایت	۲.۱
۴۷ شغلی	
۴۷ ضرایب استاندارد، انحراف استاندارد و مقدار بحرانی متغیرهای پژوهش	۳.۱
۴۸ اثرات مستقیم، اثرات غیرمستقیم و اثر کل مدل تحقیق	۴.۱
۷۳ برآورد پارامترها و خطای استاندارد آن‌ها	۱.۲
۷۴ آماره‌هایی برای ارزیابی مدل	۲.۲
۱۰۶ برآورد پارامتر برای مدل‌های نهایی در ۵ روش برآورد	۱.۳

فصل ۱

مدل سازی معادلات ساختاری

۱.۱ مقدمه

در علوم اجتماعی و رفتاری، ویژگی‌هایی مانند نگرش، ضریب هوشی، صفات شخصیتی، لیبرالیسم سیاسی، وضعیت اجتماعی و اقتصادی و... نمی‌توانند به‌طور مستقیم مشاهده یا اندازه‌گیری شوند و اغلب به آن‌ها متغیرهای پنهان^۱ (مکنون) می‌گویند. چندین دیدگاه متفاوت از متغیر مکنون وجود دارد (بولن^۲، ۲۰۰۲)، این نوع متغیرها می‌توانند با شاخص‌های متعددی که در معرض خطای اندازه‌گیری قرار دارند، محاسبه شوند. با توجه به خطای اندازه‌گیری^۳، روش‌های آماری متعارف از قبیل رگرسیون^۴، تحلیل واریانس دومتغیری^۵ و تحلیل واریانس چندمتغیری^۶ نمی‌توانند به‌طور مستقیم برای تجزیه و تحلیل روابط بین این ویژگی‌ها مورد استفاده قرار گیرند. با اختلاف خطاهای اندازه‌گیری از صفت‌ها، مدل‌سازی معادلات ساختاری^۷ SEM یک روش مستقیم برای مدل‌سازی متغیرهای پنهان فراهم می‌کند.

^۱ Latent variables

^۲ Bollen

^۳ Measurement error

^۴ Regression

^۵ ANOVA

^۶ MANOVA

^۷ Structural Equation Modeling

روش SEM در حیطه‌ی رفتار هیجانی از سال ۱۹۷۰ تحولاتی داشته است و در حال حاضر نیز به‌طور گسترده‌ای استفاده می‌شود (هرشبرگر^۸، ۲۰۰۳؛ مک کالوم و آستین^۹، ۲۰۰۰). ایده‌ی شاخص‌های متعدد از یک متغیر پنهان، از روش تحلیل عاملی می‌باشد (اسپییرمن^{۱۰}، ۱۹۰۴؛ تورستون^{۱۱}، ۱۹۷۴؛ لاولی و مکس ول^{۱۲}، ۱۹۷۱). SEM اغلب به‌عنوان یک فرمت از تحلیل عاملی، در ادبیات روان محسوب می‌شود. روش SEM چندین مدل آماری پر کاربرد در رشته‌های مختلف را پوشش می‌دهد. چند مدل خاص را قبل از معرفی ساختار کلی میانگین و کواریانس ارائه می‌کنیم. در این فصل به تفصیل در مورد تحلیل عاملی و انواع آن به‌خصوص تحلیل عاملی تاییدی، مدل‌های رگرسیون، برآورد پارامترها، تحلیل مسیر، مدل سازی معادلات ساختاری و شاخص‌های برازش با دو مثال کاربردی در شاخه مدیریت و روانشناسی می‌پردازیم. برای کسب اطلاعات بیشتر در این زمینه به قاسمی (۱۳۹۲) مراجعه شود.

۲.۱ مدل‌های رگرسیون

در کتاب‌های آماری رگرسیون به‌عنوان روش تعیین و تحلیل روابط نادقیق بین متغیرهای آماری تعریف می‌شود. بنابراین در رگرسیون، تعیین روابط بین متغیرها و تحلیل روابط به‌دست آمده مورد توجه قرار می‌گیرد. کلمه رگرسیون در لغت به معنی برگشت است. دلیل این نام‌گذاری به مطالعات گالتون^{۱۳} (۱۸۸۶) در مورد رابطه بین قد پدر و قد پسر مربوط می‌شود. به‌عنوان مثالی از رگرسیون، فرض کنید می‌خواهیم بدانیم که آیا مصرف سرانه سیگار با متغیرهایی نظیر درآمد سرانه، سطح تحصیلات، سهم دستمزد از درآمد کل و متوسط قیمت سیگار رابطه دارد یا خیر. یا ممکن است بخواهیم چگونگی اثر تغییرات یک متغیر بر متغیر دیگر را بررسی کنیم. رابطه بین مصرف سیگار با متغیرهای ذکر شده به شکل یک معادله یا الگویی است که متغیر وابسته (مصرف سیگار) را به یک یا چند متغیر مستقل یا پیش‌بین مربوط می‌کند. در رگرسیون خطی متغیر پاسخ، پیوسته است ولی متغیرهای پیش‌بین می‌توانند پیوسته، گسسته و یا رسته‌ای باشند. در مثال مصرف سیگار، متغیر وابسته بر حسب تعداد بسته‌های سیگار فروخته شده در یک استان بر مبنای سرانه در طول یک سال محاسبه می‌شود و متغیرهای پیش‌بین، متغیرهای اقتصادی-اجتماعی ذکر شده در بالا می‌باشند.

در یک معادله رگرسیونی، متغیر پاسخ را با Y و متغیرهای پیش‌بین را با X_1, X_2, \dots, X_p نشان می‌دهند. متغیرهای پیش‌بین را متغیرهای مستقل نیز می‌نامند، اما به‌کار بردن این اصطلاح ممکن است تا حدی گمراه‌کننده باشد، زیرا همیشه X_i ها، مستقل از یکدیگر نیستند

^۸Hershberger

^۹MacCallum and Austin

^{۱۰}Spearman

^{۱۱}Thurstone

^{۱۲}Lawley and Maxwell

^{۱۳}Francis Galton

و معمولا از Y مستقل نمی‌باشند. متغیر پاسخ نیز گاهی متغیر وابسته نامیده می‌شود. برای اطلاعات بیشتر، به نیرومند (۱۳۸۷) مراجعه شود.

۳.۱ رگرسیون خطی چندگانه

مدل رگرسیون خطی چندگانه عموماً به صورت زیر می‌باشد

$$Y_i = \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_p X_{ip} + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (1.1)$$

که در آن Y_i متغیر وابسته i -ام و $\beta_1, \beta_2, \dots, \beta_p$ ضرایب رگرسیونی و ε_i ها خطاهای تصادفی می‌باشد. هم‌چنین فرض می‌کنیم

$$E(\varepsilon_i) = 0, \quad Var(\varepsilon_i) = \sigma^2, \quad Cov(\varepsilon_i, \varepsilon_j) = 0, \quad i \neq j.$$

برای برآورد ضرایب رگرسیون، معمولا روش کمترین توان‌های دوم را به کار می‌گیرند.

۱.۳.۱ برآورد ضرایب رگرسیون

باید توجه داشت که برای پیدا کردن برآورد کمترین توان‌های دوم ضرایب رگرسیون، لازم نیست که فرض شود بردار خطای رگرسیون از توزیع نرمال پیروی می‌کند. مدل رگرسیون خطی چندگانه (۱.۱) را می‌توان به صورت ماتریسی زیر نیز نشان داد.

$$Y = X\beta + \varepsilon \quad (2.1)$$

که در آن

$$Y = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix}, \quad X = \begin{pmatrix} X_{11} & X_{12} & \dots & X_{1p} \\ X_{21} & X_{22} & \dots & X_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ X_{n1} & X_{n2} & \dots & X_{np} \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{pmatrix}, \quad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

برآورد کمترین توان‌های دوم β از روش کمینه کردن مجموع توان‌های دوم خطا به صورت زیر به دست می‌آید

$$\begin{aligned} \hat{\beta} &= \underset{\beta}{\operatorname{argmin}} [(Y - X\beta)^T (Y - X\beta)] \\ &= (X^T X)^{-1} X^T Y \end{aligned}$$

به طوری که $\hat{\beta} = (\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_p)^T$ یک بردار p بعدی از برآوردهای ضرایب رگرسیون است.

واضح است که برآوردگر کمترین توان های دوم برای β ناریب است، زیرا

$$\begin{aligned} E(\hat{\beta}) &= E(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y} \\ &= (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T E(\mathbf{Y}) \\ &= (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{X} \beta \\ &= \beta \end{aligned}$$

از طرفی

$$\begin{aligned} Var(\hat{\beta}) &= Var((\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}) \\ &= (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T Var(\mathbf{Y}) \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \\ &= (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T (\sigma^2 \mathbf{I}) \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \\ &= \sigma^2 (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \\ &= \sigma^2 (\mathbf{X}^T \mathbf{X})^{-1} \end{aligned}$$

از آن جایی که برآورد σ^2 از اهمیت بسیاری برخوردار است، در قضیه زیر که برگرفته از مونت گومری (۲۰۱۲) می باشد، برآوردگر ناریب σ^2 را ارائه می دهیم.

قضیه ۱.۳.۱. تحت مفروضات مدل رگرسیون خطی چندگانه (۱.۱) برآوردگر ناریب پارامتر σ^2 عبارتست از

$$S^2 = \frac{1}{n-p} \mathbf{Y}^T (\mathbf{I}_n - \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T) \mathbf{Y} = \frac{1}{n-p} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2.$$

در اغلب مسائل کاربردی واریانس مولفه خطا $\sigma^2 \mathbf{I}$ نبوده بلکه بین مولفه های خطا همبستگی وجود دارد. در این گونه مسایل واریانس خطا به صورت $\sigma^2 \mathbf{V}$ است که در آن \mathbf{V} یک ماتریس متقارن در اندازه n است.

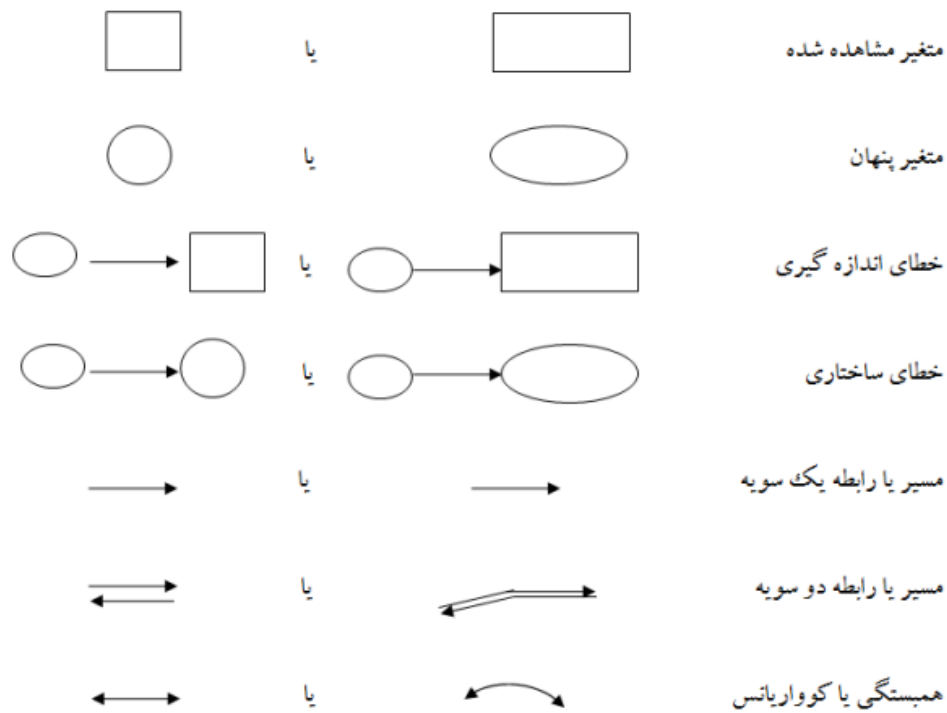
در این حالت فرضیات مدل (۱.۱) به صورت $E(\varepsilon) = 0$ و $Var(\varepsilon) = \sigma^2 \mathbf{V}$ ، که در آن \mathbf{V} یک ماتریس معین مثبت است، تغییر می کند. در این حالت محک کمترین توان های دوم به صورت $(\mathbf{Y} - \mathbf{X}\beta)^T \mathbf{V}^{-1} (\mathbf{Y} - \mathbf{X}\beta)$ می باشد. برآوردگر کمترین توان های دوم $\hat{\beta}$ به صورت زیر حاصل می شود.

$$\begin{aligned} \hat{\beta} &= \underset{\beta}{\operatorname{argmin}} [(\mathbf{Y} - \mathbf{X}\beta)^T \mathbf{V}^{-1} (\mathbf{Y} - \mathbf{X}\beta)] \\ &= (\mathbf{X}^T \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{V}^{-1} \mathbf{Y} \end{aligned}$$

۴.۱ تحلیل عاملی

بردار تصادفی $X = (X_1, X_2, \dots, X_p)^T$ از متغیرهای تصادفی را در نظر بگیرید در تحلیل‌های چندمتغیره، بزرگ بودن بُعد بردار تصادفی X که به مشقات چندبعدی^{۱۴} معروف است، اغلب در به دست آوردن روش‌های آماری مناسب برای تحلیل نمونه تصادفی موجب مشکلاتی می‌گردد. حال می‌خواهیم با از دست دادن حداقل اطلاعات، بُعد متغیرها را تا حد قابل ملاحظه‌ای تقلیل دهیم. این تفکر از آن جا ناشی می‌گردد که در مراحل اولیه تحقیق، توجه به سوی متغیرهایی متمرکز است که از یک مشاهده به مشاهده دیگر بیشترین تغییرات را نشان می‌دهند. متغیرهایی که از یک مشاهده به مشاهده دیگر زیاد عوض نمی‌شوند را می‌توان به عنوان ثابت در نظر گرفت. با کنار گذاشتن متغیرهایی با واریانس پائین و توجه به متغیرهایی با واریانس بالا، می‌توانیم به راحتی مساله خود را در یک زیر فضایی با بُعد کمتر مورد مطالعه قرار دهیم. در این راستا یکی از تکنیک‌های چندمتغیری تحلیل مولفه اصلی است. هدف اصلی این تکنیک کاهش بعد (کاهش تعداد متغیرها) یک مجموعه داده چندمتغیره است تا آن جایی که تا حد ممکن تغییرات متغیرهای اولیه در مجموعه داده را توضیح دهد. این هدف به وسیله تبدیل متغیرهای اولیه به یک مجموعه جدید از متغیرهای ناهمبسته با نام مولفه‌های اصلی به دست می‌آید که ترکیبات خطی از متغیرهای اولیه هستند و طوری مرتب شده‌اند که چند مولفه‌ی اول، بیشترین تغییرپذیری در متغیرهای اصلی را محاسبه می‌کنند. لازم به ذکر است که کاهش بعد، که اغلب به وسیله‌ی تحلیل مولفه‌ی اصلی انجام می‌گیرد تنها وقتی ممکن است که، متغیرهای اولیه همبسته باشند. قبل از ادامه‌ی مطالب در اینجا نمادها (شکل‌های) مورد استفاده در مدل‌سازی معادلات ساختاری را به صورت زیر نمایش می‌دهیم

^{۱۴} Curse of dimensionality



شکل ۱.۱: شکل های مورد استفاده در مدل سازی معادلات ساختاری

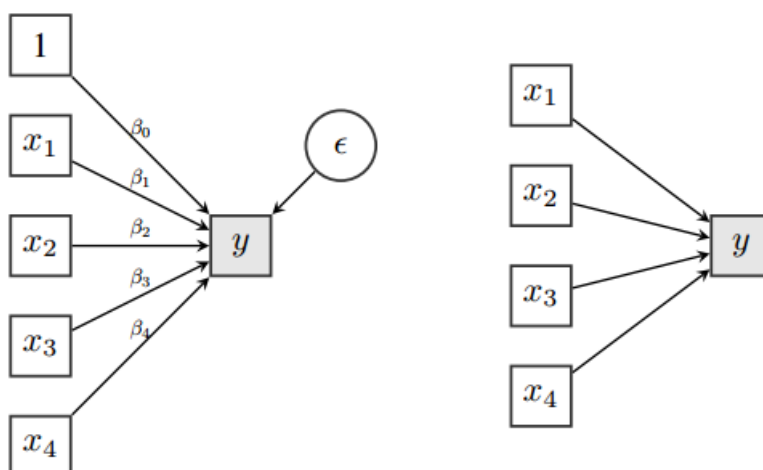
شکل ۱.۱ یک نمونه از شکل های مورد استفاده در مدل سازی معادلات ساختاری را نشان می دهد که این نمادها در رسم گراف بسیار به ما کمک می کنند.

حال مدل رگرسیون چندگانه به صورت

$$Y = X\beta + \varepsilon \quad (3.1)$$

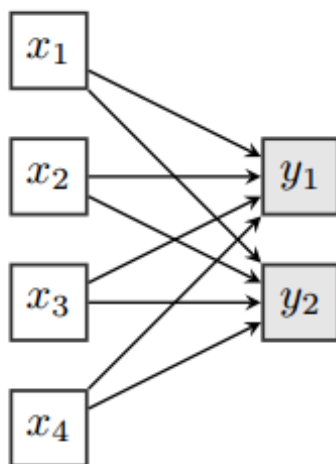
را در نظر بگیرید که در آن بردار n -تایی پاسخ، $Y = (Y_1, Y_2, \dots, Y_n)^T$ بردار $n \times p$ ماتریس طرح غیر تصادفی $X_i \in R^p$ ، i -امین متغیر پیش بین، $\beta = (\beta_1, \beta_2, \dots, \beta_p)^T$ بردار ضرایب رگرسیونی و $\varepsilon = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)^T$ بردار n -تایی شامل مولفه های خطا می باشد. سوالی که مطرح می شود این است که اگر علاقه مند باشیم این کاهش بعد در متغیرهای رگرسیونی را در قالب این مدل انجام دهیم چه طور می توانیم از روش مولفه اصلی استفاده کنیم؟ برای پاسخ به این سوال می توانیم روش مولفه های اصلی را در تحلیل عاملی به کار ببریم. دقت کنید که در مدل (۳.۱) متغیر پاسخ y روی متغیر پیش بین x رگرسیون می شود به طور خلاصه، در تحلیل عاملی عکس این مطلب اتفاق می افتد. در ادامه توضیح بیشتری در خصوص روش کار در تحلیل عاملی داده می شود.

شکل ۲.۱ یک نمونه از مدل رگرسیون چندگانه را نشان می دهد که در آن فقط یک متغیر پاسخ وجود دارد.



$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \epsilon_i \quad (i = 1, 2, \dots, n)$$

شکل ۲.۱: نمودار رگرسیون چندگانه



شکل ۳.۱: نمودار رگرسیون چندمتغیره

شکل ۳.۱ یک نمونه از مدل رگرسیون چندمتغیره را نشان می‌دهد که بیش از یک متغیر پاسخ در آن دیده می‌شود.

۱.۴.۱ انواع تحلیل عاملی

در بسیاری از حوزه‌های روان‌شناسی و رشته‌های دیگر در علوم رفتاری، اغلب این امکان وجود ندارد که به‌طور مستقیم مفاهیم ابتدایی مورد نیاز را اندازه بگیریم. به‌عنوان مثال می‌توان به هوش^{۱۵} و طبقه اجتماعی^{۱۶} اشاره کرد. در چنین مواردی محقق مجبور است که مفاهیم را به‌صورت غیرمستقیم بررسی کند یعنی اطلاعاتی از متغیرها که قابل اندازه‌گیری بوده و یا به‌صورت مستقیم مشاهده می‌شوند را جمع‌آوری کند. به‌طور مثال، روان‌شناسی که به موضوع هوش فردی علاقه‌مند است ممکن است مقادیر (نمرات) امتحانات درس‌های متفاوت، با توجه به این نکته که این نمرات روی هوش تاثیرگذار هستند، را ثبت کند. البته نباید فراموش کنیم که خطاهای تصادفی نیز وجود دارند. روان‌شناسی که روی طبقه‌ی اجتماعی مردم کار می‌کند ممکن است سوالاتی درباره‌ی حرفه‌ی افراد، تحصیلات، مالکیت خانه و غیره مطرح کند (دوباره فرض بر این است که این‌ها دقیقاً همان مفهوم مورد علاقه‌ی او را انعکاس می‌دهند). هر دوی هوش و طبقه اجتماعی عموماً به متغیرهای پنهان معروف هستند. روش تحلیلی که عموماً برای کشف روابط بین متغیرهای پنهان مفروض و متغیرهای آشکار استفاده می‌شود، تحلیل عاملی نامیده می‌شود.

روشی که مدل بر مبنای آن پایه‌ریزی می‌شود در اصل رگرسیون چندگانه است با این تفاوت که در تحلیل عاملی متغیرهای آشکار روی متغیرهای پنهان غیر قابل مشاهده (اغلب در تحلیل عاملی عامل‌های مشترک^{۱۷} یا به‌طور خلاصه عامل نامیده می‌شوند) رگرسیون می‌شوند، به‌طوری که برآورد مستقیم ضرایب رگرسیون متناظر (بار عاملی^{۱۸}) امکان‌پذیر نیست. دقت داشته باشید که تحلیل عاملی در دو نوع مجزا می‌تواند انجام شود

- تحلیل عاملی اکتشافی^{۱۹}: بررسی روابط بین متغیرهای آشکار و عامل‌ها بدون در نظر گرفتن هیچ فرضی در مورد وابسته بودن متغیرهای آشکار به عامل‌ها می‌باشد.

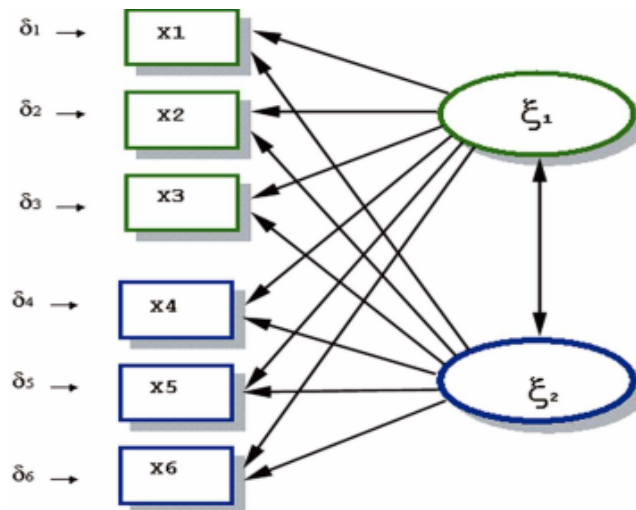
^{۱۵}Intelligence

^{۱۶}Social class

^{۱۷}Common factor

^{۱۸}Factor loading

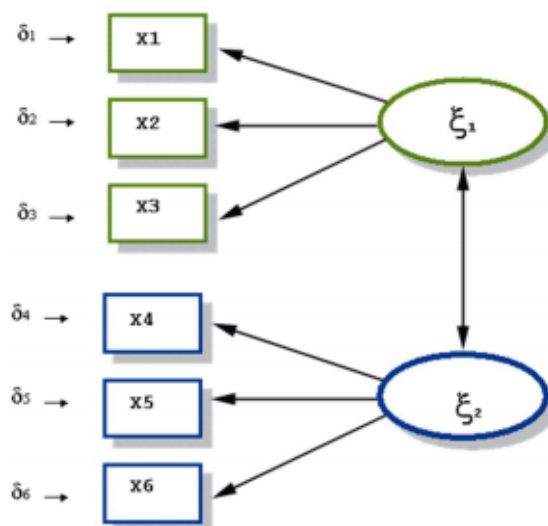
^{۱۹}Exploratory factor analysis



شکل ۴.۱: نمودار تحلیل عاملی اکتشافی

شکل ۴.۱ یک نمونه از مدل تحلیل عاملی اکتشافی را نشان می‌دهد.

- تحلیل عاملی تاییدی^{۲۰}: بررسی روابط با در نظر گرفتن وابستگی می‌باشد. به عبارت دقیق‌تر، آزمودن این که آیا یک مدل عاملی مشخص برازشی مناسب برای کواریانس‌ها یا همبستگی‌های بین متغیرهای آشکار تهیه می‌کند یا خیر.



شکل ۵.۱: نمودار تحلیل عاملی تاییدی

^{۲۰}Confirmatory factor analysis

شکل ۵.۱ یک نمونه از مدل تحلیل عاملی تاییدی را نشان می دهد.

مثال ساده از مدل تحلیل عاملی

فرض کنید ماتریس همبستگی نمونه تعدادی نمره امتحان دانش آموزان در سه موضوع زبان و ادبیات (x_1)، فرانسه (x_2) و انگلیسی (x_3) به صورت زیر باشد

$$R = \begin{matrix} x_1 & \begin{pmatrix} 1/00 \\ 0/83 & 1/00 \\ 0/78 & 0/67 & 1/00 \end{pmatrix} \\ x_2 & \\ x_3 & \end{matrix}$$

اگر فرض کنیم تنها یک عامل داریم، آن گاه مدل تک عاملی^{۲۱} به صورت زیر مشخص می شود

$$x_1 = \lambda_1 f + u_1,$$

$$x_2 = \lambda_2 f + u_2,$$

$$x_3 = \lambda_3 f + u_3$$

بدیهی است که این مدل در اصل شامل رگرسیون های خطی ساده از هر متغیر مشاهده شده روی تک عامل مشترک می باشد. در این مثال متغیر پنهان مربوطه یا عامل مشترک، f ، ممکن است نشان دهنده هوش یا توانایی عقلانی عمومی، باشد. پارامترهای λ_1 ، λ_2 و λ_3 که در واقع ضرایب رگرسیون هستند معروف به بارهای (محموله های) عاملی می باشند و عبارات u_1 ، u_2 و u_3 جملات خطای تصادفی را نشان می دهند.

k- عاملی^{۲۲}

اساس تحلیل عاملی، مدل رگرسیونی است که متغیرهای آشکار را به یک مجموعه از متغیرهای پنهان (غیرقابل مشاهده) مرتبط می سازد. در واقع، در مدل فرض می شود که روابط مشاهده شده بین متغیرهای آشکار (که به وسیله ی کواریانس ها یا همبستگی های آنها اندازه گیری می شود) نتیجه ای از روابط بین این متغیرها با متغیرهای پنهان می باشد (اگر کواریانس ها یا همبستگی های متغیرهای آشکار، مرکزی شده باشند، می توان در توصیف ریاضی روش داده شده فرض کرد که متغیرهای آشکار میانگین صفر دارند).

برای شروع فرض می کنیم که یک مجموعه از متغیرهای آشکار یا مشاهده شده داریم $X^T = (X_1, X_2, \dots, X_q)$ که فرض شده که با k متغیر پنهان غیر مشاهده شده یا عامل های مشترک f_1, \dots, f_k که در آن $k < q$ به وسیله ی مدل رگرسیون زیر مرتبط می باشند.

$$X_1 = \mu_1 + \lambda_{11}f_1 + \lambda_{12}f_2 + \dots + \lambda_{1k}f_k + u_1$$

$$X_2 = \mu_2 + \lambda_{21}f_1 + \lambda_{22}f_2 + \dots + \lambda_{2k}f_k + u_2$$

⋮

$$X_q = \mu_q + \lambda_{q1}f_1 + \lambda_{q2}f_2 + \dots + \lambda_{qk}f_k + u_q$$

^{۲۱}Single-factor

^{۲۲}The K-factor analysis model

که در آن X_i ها مشاهدات، μ_i ها میانگین، λ_{ij} بارهای عاملی، f_i عامل ها و u_i خطاها را در مدل فوق نشان می دهند. در تحلیل عاملی، ضرایب رگرسیون λ_j با عنوان بار عاملی، شناخته می شوند. معادلات رگرسیون در بالا ممکن است به صورت مختصر زیر نوشته شوند

$$X = \mu + \Lambda f + u$$

که در آن $E(u) = 0$ ، $E(f) = 0$ ، $E(X) = \mu$

$$\Lambda = \begin{pmatrix} \lambda_{11} & \lambda_{12} & \dots & \lambda_{1k} \\ \vdots & & & \vdots \\ \lambda_{q1} & \lambda_{q2} & \dots & \lambda_{qk} \end{pmatrix}, \quad f = \begin{pmatrix} f_1 \\ \vdots \\ f_k \end{pmatrix}, \quad u = \begin{pmatrix} u_1 \\ \vdots \\ u_q \end{pmatrix}, \quad \mu = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_q \end{pmatrix}$$

f بردار متغیرهای پنهان و Λ ماتریس بارهای عاملی و u بردار خطای اندازه گیری نامیده می شود.

فرض کنید که عبارات آشفته تصادفی u_1, u_2, \dots, u_q با یکدیگر و با عامل های f_1, f_2, \dots, f_k ناهمبسته باشند (عناصر u با هر X_i مشخص می شوند و از این رو به طور کلی در این مدل با عنوان تغییرات مشخص^{۲۳} شناخته می شوند) این دو فرض نتیجه می دهند که به شرط معلوم بودن مقادیر عامل های مشترک، متغیرهای آشکار متغیرهای مستقل هستند یعنی همبستگی های متغیرهای مشاهده شده، از روابط آن ها با عامل های مشترک نشات می گیرد. از آن جایی که عامل ها مشاهده نشده اند، می توان مکان و مقیاس آن ها را به صورت دلخواه ثابت اختیار کرد. بنابراین فرض می کنیم که استاندارد شده اند به عبارتی دارای میانگین صفر و انحراف معیار یک هستند. به علاوه فرض می کنیم که عامل ها با یکدیگر ناهمبسته اند. در این مورد بارهای (محموله های) عاملی، همبستگی های متغیرهای آشکار و عامل ها می باشند. با این فرضیات اضافی در مورد عامل ها، مدل تحلیل عاملی نتیجه می دهد که واریانس متغیر X_i یعنی σ_i^2 به صورت زیر است

$$\sigma_i^2 = \sum_{j=1}^k \lambda_{ij}^2 + \psi_i$$

که ψ_i واریانس u_i می باشد. در نتیجه مشاهده کردیم که مدل تحلیل عاملی نتیجه می دهد که واریانس هر متغیر مشاهده شده می تواند به دو قسمت تقسیم شود: اول h_i^2 که به صورت زیر می باشد

$$h_i^2 = \sum_{j=1}^k \lambda_{ij}^2$$

و با عنوان واریانس اشتراکی^{۲۴} (همه داشت یا مشترک) متغیر معروف می باشد و واریانس سهمیم شده با متغیرهای دیگر با استفاده از عامل های معمول را نمایش می دهد. قسمت دوم،

^{۲۳} Specific variates

^{۲۴} Communality

ψ_i واریانس یکتا یا مشخص^{۲۵} (خاص) نامیده می شود و مربوط به تغییراتی در X_i است که با متغیرهای دیگر سهمیم نشده است. به علاوه مدل عامل، عبارت زیر برای کواریانس متغیرهای X_i و X_j را نتیجه می دهد

$$\sigma_{ij} = \sum_{l=1}^k \lambda_{il} \lambda_{jl}$$

مشاهده می کنیم که کواریانس ها از تغییرات مشخص در هر روش مستقل نیستند، تنها عامل های مشترک برای محاسبه ی روابط بین متغیرهای آشکار مورد هدف می باشند. نتایج بالا نشان می دهد که مدل تحلیل k عاملی نتیجه می دهد که ماتریس کواریانس جامعه Σ از متغیرهای مشاهده شده به صورت زیر می باشد

$$\Sigma = \Lambda \Lambda^T + \Psi$$

که در آن $Cov(u) = \Psi$ ، $Cov(X) = \Sigma$ به طوری که $\Psi = diag(\psi_i)$. حالت عکس نیز برقرار است، اگر Σ بتواند به شکل داده شده در بالا تجزیه شود آن گاه مدل k- عاملی برای X برقرار است. به عبارت دقیق تر چنان چه بخواهیم بررسی کنیم که آیا مدل k- عاملی برای مشاهدات داده شده مناسب است کافی است ساختار کواریانس آن ها را بررسی کنیم.

در عمل، Σ به وسیله ی ماتریس کواریانس نمونه ای (S) برآورد می شود که بنابراین نیاز داریم برآوردهای Λ و Ψ را به دست آوریم، به طوری که ماتریس کواریانس مشاهده شده، در صورت ساختاری مورد نیاز برای مدل عاملی ($\Sigma = \Lambda \Lambda^T + \Psi$) صدق کند. هم چنین نیاز خواهیم داشت که مقدار k، تعداد عامل ها برای تهیه یک برازش کافی برای S را تعیین کنیم.

۲.۴.۱ تحلیل عاملی تاییدی

تحلیل اکتشافی مشخص می کند که کدام متغیرهای مشاهده شده با عامل مشترک به شدت همبسته بوده و چه تعداد عامل مشترک برای ارائه توصیفی کافی از داده ها لازم است. در یک تحلیل عاملی اکتشافی هیچ محدودیتی روی متغیرهایی که روی عامل ها بارگذاری می شوند قرار داده نشده است. در مدل های تحلیل عاملی تاییدی، فرض می شود که متغیرهای آشکار بخصوصی روی عامل خاصی بارگذاری می شوند در حالی که دیگر متغیرهای آشکار این محدودیت را دارند که روی چند عامل بارهای صفر دارند (یا به عبارتی بارگذاری نمی شوند) بنابراین این مدل مبتنی بر اطلاعات پیش تجربی درباره ساختار داده هاست که می تواند در قالب فرضیه ای، یک طرح طبقه بندی در انطباق با ویژگی های عینی شکل و محتوا یا دانش حاصل از مطالعات قبلی درباره داده ها می باشد. بنابراین یک مدل تحلیل عاملی تاییدی ممکن است بر اساس نتایج یک تحلیل عاملی اکتشافی باشد که در آن محقق تمایل دارد مدلی مشخص را برای یک

^{۲۵}Specific

مجموعه‌ی جدید از داده‌های مشابه، قطعی فرض کند بدین معنی که بارهای چند متغیر آشکار روی چند عامل را صفر در نظر می‌گیرد چون آن بارها در تحلیل اکتشافی کوچک بودند. تفاوت عمده تحلیل عاملی تاییدی با تحلیل عاملی اکتشافی در این است که در این حالت ممکن است چند جفت از عامل‌ها همبسته در نظر گرفته شوند (با توجه به اطلاعات به دست آمده قبلی). بنابراین در این مدل فرض می‌شود بین عوامل پنهان همبستگی وجود داشته باشد که می‌توان این همبستگی را در مدل بندی وارد نمود. برای این منظور فرض کنید

$$Cov(\mathbf{f}) = \Phi = (\varphi_{ij})$$

در این صورت مدل کواریانس X به صورت زیر است.

$$\Sigma = \Lambda\Phi\Lambda^T + \Psi$$

حال اگر فرض کنید θ نشان دهنده‌ی مجهولات در Φ ، Λ و Ψ باشد معمولاً مدل کواریانس فوق را به صورت زیر نشان می‌دهند

$$\Sigma(\theta) = \Lambda\Phi\Lambda^T + \Psi$$

در تحلیل عاملی هدف برآورد θ و به تبع آن برآورد $\Sigma(\theta)$ است.

مدل تحلیل عاملی تاییدی زیر مجموعه‌ای از روش کلی تر برای مدل بندی متغیرهای پنهان معروف به مدل بندی معادلات ساختاری یا مدل بندی ساختار کواریانس است. مدل سازی معادله ساختاری به طور معمول ترکیبی از مدل های اندازه گیری^{۲۶} و مدل های ساختاری^{۲۷} است. بر مبنای مدل های اندازه گیری، محقق تعریف می‌کند که کدام متغیرهای مشاهده شده یا معرف های اندازه گیرنده‌ی کدام متغیرهای پنهان هستند و بر پایه مدل ساختاری مشخص می‌شود که کدام متغیرها با یکدیگر همبسته‌اند. به این ترتیب با بهره گیری از این مدل ها می‌توان به طور همزمان به ارزیابی کیفیت سنجش متغیرها و مقبولیت اثرات مستقیم و غیرمستقیم و همچنین تعامل های تعریف شده میان متغیرها پرداخت. با فرض اینکه $\eta \sim N(\circ, \psi)$ و $\varepsilon \sim N(\circ, \Theta)$ و $cov(\varepsilon, \eta) = \circ$ یعنی η و ε ناهمبسته‌اند. مدل تحلیل عاملی تاییدی در فرمت ماتریس به صورت زیر می‌باشد.

$$\begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_p \end{pmatrix}_{p \times 1} = \begin{pmatrix} \lambda_{11} & \lambda_{12} & \dots & \lambda_{1m} \\ \lambda_{21} & \lambda_{22} & \dots & \lambda_{2m} \\ \vdots & \vdots & & \vdots \\ \lambda_{p1} & \lambda_{p2} & \dots & \lambda_{pm} \end{pmatrix}_{p \times m} \begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_p \end{pmatrix}_{m \times 1} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_p \end{pmatrix}_{p \times 1}$$

$$\mathbf{y} = \Lambda\boldsymbol{\eta} + \boldsymbol{\varepsilon}$$

^{۲۶} Measurement models

^{۲۷} Structure models

۳.۴.۱ برآورد پارامترها به روش ماکزیمم درستنمایی

فرض کنید داده‌ها دارای توزیع نرمال چندمتغیره باشند و مدل عاملی برای آن‌ها برقرار باشد (یعنی ساختار کواریانس به صورت $\Sigma = \Lambda\Phi\Lambda^T + \Psi$ باشد). در این صورت تابع چگالی احتمال بردار تصادفی q بعدی مشاهده x_i به صورت زیر است

$$f(x_i; \mu, \Lambda, \Phi, \Psi) = (\gamma\pi)^{-\frac{q}{\gamma}} |\Lambda\Phi\Lambda^T + \Psi|^{-\frac{1}{\gamma}} \exp\left(-\frac{1}{\gamma}(x_i - \mu)^T(\Lambda\Phi\Lambda^T + \Psi)^{-1}(x_i - \mu)\right)$$

بر اساس یک نمونه تصادفی n تایی، تابع درستنمایی برابر است با

$$L(\mu, \Lambda, \Phi, \Psi) = (\gamma\pi)^{-\frac{nq}{\gamma}} |\Lambda\Phi\Lambda^T + \Psi|^{-\frac{n}{\gamma}} \exp\left(-\frac{1}{\gamma} \sum_{i=1}^n (x_i - \mu)^T(\Lambda\Phi\Lambda^T + \Psi)^{-1}(x_i - \mu)\right)$$

با توجه به این که

$$\begin{aligned} & \sum_{i=1}^n (x_i - \mu)^T(\Lambda\Phi\Lambda^T + \Psi)^{-1}(x_i - \mu) \\ &= \sum_{i=1}^n (x_i - \mu - \bar{x} + \bar{x})^T(\Lambda\Phi\Lambda^T + \Psi)^{-1}(x_i - \mu - \bar{x} + \bar{x}) \\ &= \sum_{i=1}^n (x_i - \bar{x})^T(\Lambda\Phi\Lambda^T + \Psi)^{-1}(x_i - \bar{x}) + n(\bar{x} - \mu)^T(\Lambda\Phi\Lambda^T + \Psi)^{-1}(\bar{x} - \mu) \\ &= ntr((\Lambda\Phi\Lambda^T + \Psi)^{-1}S) + n(\bar{x} - \mu)^T(\Lambda\Phi\Lambda^T + \Psi)^{-1}(\bar{x} - \mu) \end{aligned}$$

لگاریتم تابع درستنمایی در نقطه $\hat{\mu} = \bar{x}_i$ برابر است با

$$l(\Phi, \Lambda, \Psi) = -\frac{nq}{\gamma} \ln(\gamma\pi) - \frac{n}{\gamma} \ln |\Lambda\Phi\Lambda^T + \Psi| - \frac{n}{\gamma} tr[(\Lambda\Phi\Lambda^T + \Psi)^{-1}S]$$

با اضافه و کم کردن $\ln|S|$ (تابعی از مشاهدات)، لگاریتم تابع درستنمایی می‌تواند به صورت $-\frac{1}{\gamma}nF$ به علاوه یک تابع از مشاهدات نشان داده شود که F به صورت زیر است

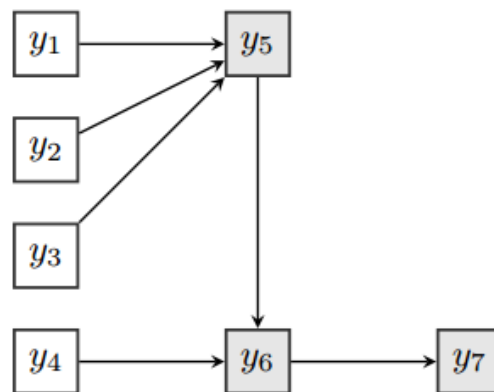
$$\begin{aligned} F &= \ln |\Lambda\Phi\Lambda^T + \Psi| + trace(S(\Lambda\Phi\Lambda^T + \Psi)^{-1}) - \ln|S| - q \\ &= \ln |\Sigma(\theta)| + tr(S\Sigma^{-1}(\theta)) - \ln|S| - q \end{aligned}$$

تابع F مقدار صفر را می‌گیرد اگر $\Lambda\Phi\Lambda^T + \Psi$ مساوی با S باشد و در غیر این صورت مقادیر بزرگتر از صفر را می‌گیرد. برآوردهای بارها و واریانس‌های مشخص با کمینه کردن F نسبت به این پارامترها به دست می‌آید.

۵.۱ تحلیل مسیر

تحلیل مسیر به عنوان روشی برای مطالعه تاثیرات مستقیم و غیرمستقیم معرفی شده است. در حقیقت تحلیل مسیر نموداری است که نشان می‌دهد چگونه یک مجموعه از متغیرهای اکتشافی روی یک متغیر تحت بررسی تاثیر می‌گذارند. مسیر^{۲۸} در مدل علی^{۲۹} نشان‌دهنده اثر یک متغیر بر متغیر دیگر است. در تحلیل مسیر^{۳۰} معمولاً مسیر را با یک فلش جهت دار یک طرفه که از متغیر پیش‌بین (برون‌زا^{۳۱}) به متغیر مربوطه وابسته (درون‌زا^{۳۲}) رسم شده است نمایش می‌دهند. نمایش p_{ij} مسیر دارای یک نمایش ریاضی است که به صورت عمومی داده می‌شود. حرف i نشان‌گر متغیر برون‌زا حرف j نشان‌گر متغیر درون‌زا است و همواره اصل $j > i$ برقرار است به عبارت دیگر i متغیر اثرپذیر و j متغیر اثرگذار بر آن است. پس مسیر فرضی $p_{۴۱}$ یعنی یک متغیر بر چهار متغیر موثر است یا این که متغیر یک، متغیر برون‌زا و متغیر چهار، متغیر درون‌زا است.

مدل رگرسیون (۳.۱) را در نظر بگیرید نمودار مسیر این مدل عبارتست از



شکل ۶.۱: نمودار مسیر

شکل ۶.۱ یک نمونه از مدل مسیر را نشان می‌دهد.

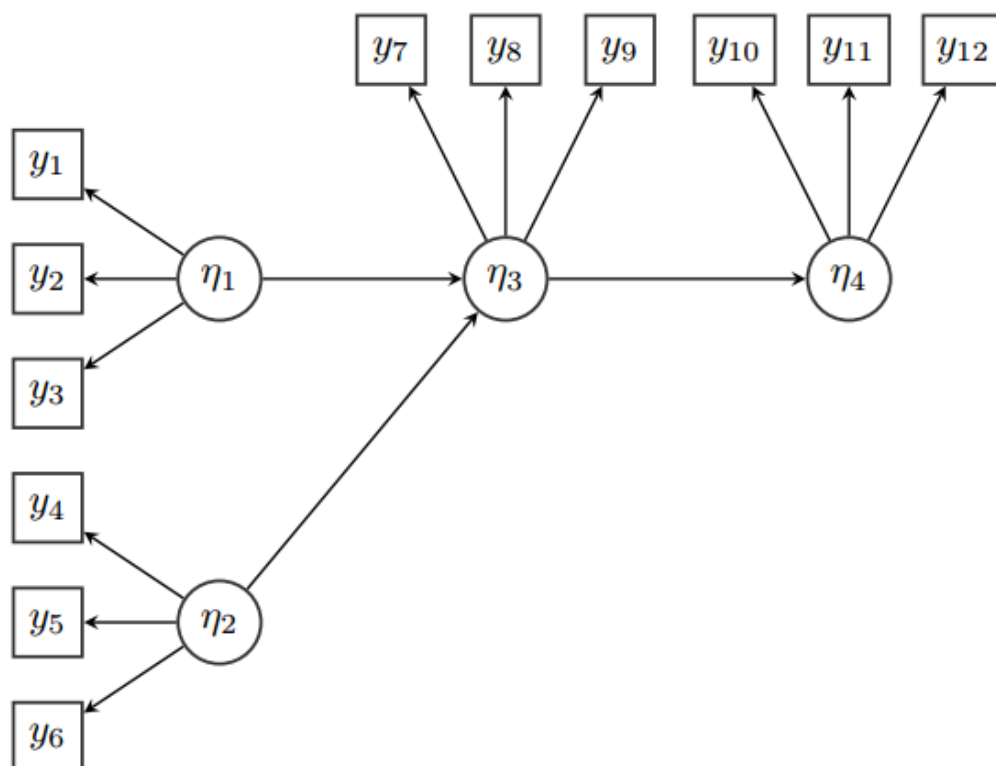
^{۲۸} path

^{۲۹} Causal model

^{۳۰} Path diagram

^{۳۱} Exogenous variable

^{۳۲} Endogenous variable



شکل ۷.۱: نمودار مسیر با متغیرهای پنهان

و شکل ۷.۱ یک نمونه از مدل مسیر با متغیرهای پنهان را نشان می‌دهد. در تحلیل مسیر یک متغیر ممکن است همزمان نقش متغیر پیش‌بین یا وابسته را ایفاء نماید. به عبارت دیگر، یک متغیر در مدل علی ممکن است نسبت به برخی متغیرها پیش‌بین و نسبت به برخی دیگر وابسته باشد. برای جلوگیری از ابهام و سردرگمی به جای پیش‌بین و وابسته از دو اصطلاح دیگر برای تعیین نوع متغیرها در روش تحلیل مسیر استفاده می‌شود.

درون‌زا (وابسته) و برون‌زا (پیش‌بین)

کلیه متغیرهای موجود در یک مدل و الگوی علی دارای دو نوع اصلی است. نوع اول متغیر برون‌زا و نوع دوم متغیر درون‌زا نام دارد. متغیر برون‌زا متغیری است که هیچ اثری از سایر متغیرهای الگو و مدل طراحی شده نمی‌پذیرد. در حقیقت مقدار متغیر برون‌زا توسط سایر متغیرهای درون مدل تعیین نمی‌شود بلکه مقدار آن در خارج مدل تعیین می‌شود.

متغیر درون‌زا (وابسته) متغیری است که از حداقل یک متغیر دیگر در مدل و الگوی طراحی شده اثر می‌پذیرد. مقدار متغیر درون‌زا توسط سایر متغیرهای درون مدل تعیین می‌شود. بنابراین بر اساس تعریف، یک متغیر نمی‌تواند همزمان هم درون‌زا و هم برون‌زا باشد. از نظر نموداری متغیر برون‌زا متغیری است که هیچ فلشی به آن وارد نمی‌شود در حالی که متغیر درون‌زا متغیری است که حداقل یک فلش به آن وارد می‌شود.

تعاریف تحلیل مسیر

- مسیر: هر خط دارای جهت را یک مسیر می‌نامند.

- ضریب مسیر^{۳۳}: ضرایب بتای استاندارد در خروجی رگرسیون را ضریب مسیر گویند.
- نمودار تحلیل مسیر: نمایش ترسیمی تحلیل مسیر را نمودار تحلیل مسیر گویند.
- ورودی نمودار مسیر: برای ترسیم مدل مفهومی و قبل از آزمون است.
- خروجی نمودار مسیر: مدل مفهومی است که آزمون شده است.
- مدل اصلاح شده^{۳۴}: مدلی است که مسیرهای غیرمعنادار حذف شده و مسیرهای جایگزین به آن اضافه شده است.

هدف تحلیل مسیر به دست آوردن برآوردهای کمی روابط علی بین مجموعه‌ای از متغیرهاست. ساختن یک مدل علی لزوماً به معنای وجود روابط علی در بین متغیرهای مدل نیست بلکه این علیت بر اساس مفروضات همبستگی و نظر و پیشینه تحقیق استوار است. تحلیل مسیر بیان می‌کند که کدام مسیر مهم‌تر و یا معنادارتر است. ضرایب مسیر براساس ضریب استاندارد شده رگرسیون محاسبه می‌شود. یک متغیر به صورت تابعی از دیگر متغیرها فرض می‌شود و مدل رگرسیونی آن ترسیم می‌شود. برای به دست آوردن برآوردهای ضرایب اصلی مسیر کافی است هر متغیر وابسته (درون‌زا) به متغیرهایی که مستقیماً تحت تأثیر آن است بازگشت داده شود. به بیان دیگر برای برآوردهای هر یک از مسیرهای مشخص شده، ضرایب استاندارد شده رگرسیون (یا ضرایب مسیر) محاسبه می‌شود. این ضرایب از طریق برقراری معادله‌های ساختاری یعنی معادله‌هایی که ساختار روابط مفروض در یک مدل را مشخص می‌سازد به دست می‌آیند. تحلیل مسیر صرفاً بر روی متغیرهای مشاهده شده^{۳۵} قابل انجام است.

۱.۵.۱ طراحی مدل مسیر

محقق بر اساس تحقیقات قبلی مشابه و مرتبط اقدام به انتخاب متغیرها و تعیین روابط علی بین آنها بر اساس منطق تحلیلی و نظری می‌نماید. خروجی این مرحله ممکن است مجموعه‌ای از فرضیه‌های مرتبط و یکپارچه باشد که معمولاً از طریق ترسیمی و یا ریاضی بیان می‌شود.

۲.۵.۱ انواع مدل‌های تحلیل مسیر

در یک دسته‌بندی کلی می‌توان ۶ نوع مدل تحلیل مسیر در نظر گرفت. ۱- مدل متغیرهای پیش‌بین^{۳۶} همان رگرسیون چندگانه است اما بین متغیرهای پیش‌بین همبستگی ترسیم نمی‌شود. ۲- مدل همبسته^{۳۷} مدل همبسته همانند مدل متغیرهای پیش‌بین است با این تفاوت که

^{۳۳} Path coefficient

^{۳۴} Modified path diagram

^{۳۵} observed variable

^{۳۶} Independent variables model

^{۳۷} Correlated model

بین برخی متغیرهای مدل همبستگی وجود دارد. ۳- مدل‌های دارای متغیر میانجی^{۳۸} حداقل یک متغیر واسطه بین دو متغیر دیگر قرار می‌گیرد. ۴- مدل‌های دارای متغیر تعدیل‌گر^{۳۹} یک متغیر بر ارتباط بین دو متغیر دیگر اثر تعدیل‌گر دارد. ۵- مدل‌های یک طرفه^{۴۰} جهت فلش‌ها به یک سمت بوده و برگشت به عقب ندارد یعنی همه مسیرها به یک سو هستند. ۶- مدل‌های دوطرفه^{۴۱} جهت فلش‌ها و مسیرها دارای حرکت رو به عقب بوده و یک حلقه درست می‌کند.

۳.۵.۱ انواع روابط بین متغیرها در نمودار مدل مسیر

یک متغیر پیش‌بینی‌کننده ممکن است به صورت اثر مستقیم، غیرمستقیم، کاذب^{۴۲} و تحلیل نشده بر یک متغیر وابسته تأثیر بگذارد. در اثر مستقیم متغیر X مستقیماً بر روی متغیر Y تأثیرگذار است. چنانچه X بر Y از طریق متغیری دیگر تأثیر بگذارد اثر غیرمستقیم است مثلاً اگر X بر Z اثر مستقیم داشته باشد و Z بر Y، آن‌گاه می‌گوییم X بر Y اثر غیرمستقیم دارد. بسیاری از پژوهشگران مایلند اثر کلی یک متغیر را بر متغیر دیگر محاسبه کنند این کار از طریق جمع اثر مستقیم با مجموع آثار غیرمستقیم آن به دست می‌آید. آثار غیرمستقیم از طریق حاصل ضرب ضرایب هر مسیر محاسبه می‌شود. وقتی اثر X بر Y کاذب است که Z علت هر دو متغیر X و Y باشد. و در آخر رابطه بین دو متغیر وقتی تحلیل نشده است که هر دوی آن‌ها برون‌زا بوده و بنابراین تبیین‌پذیری بین آن‌ها توسط مدل امکان‌پذیر نباشد.

۴.۵.۱ مفروضات تحلیل مسیر

- کلاین (۱۹۸۹) توصیه می‌کند که به ازای هر پارامتر (نه متغیر) در مدل حداقل ۱۰ عدد نمونه وجود داشته باشد. در نظر گرفتن نسبت ۲۰ نمونه برای هر پارامتر بسیار مطلوبست.
- متغیرهای مدل باید دارای مقیاس فاصله‌ای و نسبی باشند.
- بین متغیرهای پیش‌بین با متغیر وابسته رابطه خطی وجود دارد.
- جملات خطای متغیرها با یکدیگر ناهمبسته‌اند.
- داده‌ها دارای توزیع نرمال هستند.

^{۳۸} Mediated variable

^{۳۹} Moderator variable

^{۴۰} Recursive models

^{۴۱} Nonrecursive models

^{۴۲} Spurious

- همخطی چندگانه^{۴۳} وجود ندارد. در حقیقت هیچ دو متغیری با یکدیگر همبستگی بالایی ندارند.
- بتوان همبستگی مدل را به مجموع همبستگی های اثرات مستقیم و غیرمستقیم تجزیه کرد.

۵.۵.۱ اصول ترسیم نمودار مسیر

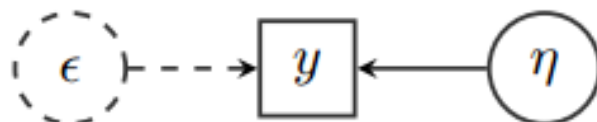
- حلقه در مدل وجود نداشته باشد.
- مسیر رفت و برگشت بین متغیرها وجود نداشته باشد.
- تعداد همبستگی های مجاز بین متغیرهای درونزا حداکثر برابر با تعداد مسیرها باشد.

۶.۱ مدل سازی معادلات ساختاری

همان طور که جلوتر به آن اشاره شده مدل معادلات ساختاری حالت کلی تری از مدل تحلیل عاملی تاییدی می باشد.

مدل اندازه گیری

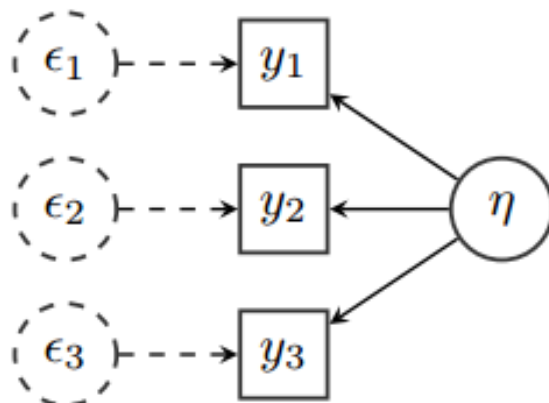
در علوم اجتماعی، مشاهده‌ی متغیرها بدون خطای اندازه‌گیری نیست. مدل اندازه‌گیری تکی این گونه نشان داده می‌شود.



شکل ۸.۱: نمودار مدل اندازه‌گیری تکی

شکل ۸.۱ نمودار مدل اندازه‌گیری تکی را نشان می‌دهد.

مدل اندازه گیری چندگانه به صورت زیر شکل ۹.۱ نشان داده می شود.



شکل ۹.۱: نمودار مدل اندازه گیری چندگانه

۱.۶.۱ الگوی ریاضی مدل معادلات ساختاری

الگوهای معادلات ساختاری شامل تعداد زیادی از تکنیک های آماری می باشند که در علوم رفتاری کاربرد فراوانی دارند. می توان این الگوها را ترکیبی از تحلیل عاملی و رگرسیون یا تحلیل مسیر نامید. این مدل ها در واقع به ساختار کواریانس بین متغیرهای مشاهده شده اشاره دارد و از همین جهت گاه مدل های ساختار کواریانس نیز خوانده می شوند. اغلب برای تعیین ساختار مدل از نمودار تحلیل مسیر استفاده می شود. الگوهای معادله ساختاری در واقع مجموعه ای از معادلات خطی را ارائه می دهند که با تعیین روابط علی بین متغیرهای پنهان سعی در کشف اثرات مستقیم و غیر مستقیم متغیرهای پنهان برونزا بر متغیرهای پنهان درونزا دارد. مدلیابی برای این اساس شامل پنج مرحله تدوین مدل، شناسایی مدل، برآورد مدل، ارزیابی مدل و اصلاح مدل می باشد. نرم افزارهای رایانه ای برای تعیین، برازش و ارزیابی الگوهای ساختاری توسط جورسکوگ و سوربوم (۱۹۶۹) بسط داده شده است و اکنون به طور وسیعی به صورت سیستم لیزرل (روابط ساختاری خطی) مورد استفاده قرار می گیرد. قبل از این که مدل ریاضی معادلات ساختاری را مطرح کنیم لازم است اشاره ای مختصر به انواع متغیرهای این مدل داشته باشیم.

۲.۶.۱ انواع متغیرها در مدل‌یابی معادلات ساختاری

- متغیرهای مشاهده پذیر^{۴۴} (آشکار): این متغیرها به صورت مستقیم قابل اندازه‌گیری هستند. که در نمودار مسیر این متغیرها را با شکل مستطیل نشان می‌دهند.
- متغیرهای پنهان^{۴۵} (مکنون^{۴۶}): این متغیرها به صورت مستقیم قابل اندازه‌گیری نیستند، بلکه از طریق یک مجموعه از شاخص‌ها یا متغیرهای مشاهده پذیر بررسی می‌شوند. که در نمودار مسیر این متغیرها را با شکل بیضی نشان می‌دهند.
- متغیرهای پیش‌بین^{۴۷} یا برون‌زا^{۴۸}: این متغیرها در مدل مفهومی^{۴۹} صرفاً دارای نقش علت هستند، یعنی فقط مسیری از آن‌ها خارج شده و نسبت به متغیرهای دیگر نقش علت را دارند، این متغیرها را تحت عنوان (کسای) نیز مطرح می‌کنند.
- متغیرهای وابسته^{۵۰} یا درون‌زا^{۵۱}: در مدل مفهومی^{۵۲} صرفاً دارای نقش معلول هستند، یعنی فقط مسیری از متغیرهای دیگر به آن‌ها ارتباط داده شده‌اند، این متغیرها نسبت به متغیرهای دیگر نقش معلول را دارند، این متغیرها را تحت عناوین (ملاک و اتا) نیز مطرح می‌کنند.
- متغیرهای پنهان میانجی^{۵۳}: در مدل مفهومی دارای دونقش علت و معلول هستند، یعنی نسبت به بعضی از متغیرها نقش علت و نسبت به بعضی دیگر نقش معلول را دارند، این متغیرها را تحت عناوین (واسطه‌ای و اتا) نیز مطرح می‌کنند.
- متغیرهای تعدیل‌کننده^{۵۴}: به صورت مستقیم بر جهت یا میزان رابطه‌ی متغیرهای پیش‌بین و وابسته می‌تواند موثر باشد. اثرات این متغیر قابل مشاهده و اندازه‌گیری می‌باشد.

متغیرهای مشاهده پذیر X: جهت بیان نشانگرهای مربوط به متغیرهای پنهان علت (که معمولاً با ξ نشان داده می‌شود) به کار برده می‌شوند و با نماد مستطیل نشان داده می‌شوند.

^{۴۴} Observe variables

^{۴۵} Unobserve variables

^{۴۶} Latent

^{۴۷} Independent variables

^{۴۸} Endogenous

^{۴۹} مدل مفهومی در حقیقت یک مدل گرافیکی است که نحوه‌ی ارتباط متغیرها در آن مشخص بوده و می‌توان کل مسئله مورد تحقیق را به صورت نموداری به مخاطب نشان داد.

^{۵۰} Dependent variables

^{۵۱} Exogenous

^{۵۲} Conceptual Model

^{۵۳} Interface latent variables

^{۵۴} Moderator variables

متغیرهای مشاهده پذیر Y : جهت بیان نشانگرهای مربوط به متغیرهای پنهان میانجی و معلول (که معمولا با η نشان داده می شود) به کار برده می شوند و با نماد مستطیل نشان داده می شوند.

فرض کنید در مطالعه ای متغیرهای علت و معلول از پیش تعیین شده ای وجود دارد. یک روش مدل بندی این متغیرها استفاده از مدل معادلات ساختاری است. در حقیقت این مدل مجموعه ای از معادلات خطی است که برای مدل بندی متغیرهای علت و معلول به کار می رود. فرض کنید متغیرهای علت و معلول، که به طور مستقیم مشاهده نمی شوند و پنهان می باشند، را به ترتیب با ξ و η نشان دهیم که η متغیر وابسته پنهان و ξ متغیر پیش بین پنهان است. هم چنین فرض کنید X و Y بردارهای قابل مشاهده و اندازه گیری باشند به طوری که رابطه های خطی زیر برقرار باشند

$$\begin{cases} X = \Lambda_x \xi + \delta \\ Y = \Lambda_y \eta + \varepsilon \end{cases} \quad (4.1)$$

به مدل (۴.۱) مدل اندازه گیری می گویند زیرا بردار داده ها $Z = (X^T, Y^T)^T$ قابل اندازه گیری می باشند.

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}_{p \times 1} = \begin{pmatrix} \lambda_{11} & \lambda_{12} & \cdots & \lambda_{1m} \\ \lambda_{21} & \lambda_{22} & \cdots & \lambda_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{p1} & \lambda_{p2} & \cdots & \lambda_{pm} \end{pmatrix}_{p \times m} \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_m \end{pmatrix}_{m \times 1} + \begin{pmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_p \end{pmatrix}_{p \times 1}$$

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{pmatrix}_{p \times 1} = \begin{pmatrix} \lambda_{11} & \lambda_{12} & \cdots & \lambda_{1m} \\ \lambda_{21} & \lambda_{22} & \cdots & \lambda_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{p1} & \lambda_{p2} & \cdots & \lambda_{pm} \end{pmatrix}_{p \times m} \begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_m \end{pmatrix}_{m \times 1} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_p \end{pmatrix}_{p \times 1}$$

حال فرض کنید داریم

$$x_1 = 1\xi + \delta_1, \quad x_2 = \lambda_2\xi + \delta_2, \quad x_3 = \lambda_3\xi + \delta_3$$

$$y_1 = 1\eta_1 + \varepsilon_1, \quad y_2 = \lambda_2\eta_1 + \varepsilon_2, \quad y_3 = \lambda_3\eta_1 + \varepsilon_3$$

$$y_4 = 1\eta_2 + \varepsilon_4, \quad y_5 = \lambda_5\eta_2 + \varepsilon_5, \quad y_6 = \lambda_6\eta_2 + \varepsilon_6$$

فرم ماتریسی مدل اندازه گیری بالا به صورت زیر می باشد

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} (\xi) + \begin{pmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \end{pmatrix}$$

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \lambda_2 & 0 \\ \lambda_3 & 0 \\ 0 & 1 \\ 0 & \lambda_5 \\ 0 & \lambda_6 \end{pmatrix} \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \end{pmatrix}$$

روابط بین متغیرهای آشکار و پنهان به وسیله بارهای عاملی (یعنی $\lambda_1 - \lambda_6$) نشان داده می شود. بارهای عاملی تعیین می کند که یک متغیر آشکار تا چه حد در سنجش متغیر پنهان توانایی دارد. بارهای عاملی به عنوان ضرایب اعتبار^{۵۵} نیز شناخته می شوند زیرا ضرب بار عاملی در نمره متغیر آشکار نشان می دهد که نمره متغیر آشکار تا چه حد معتبر است. حال فرض کنید متغیرهای پنهان وابسته $\eta = (\eta_1^*, \eta_2^*)^T$ و پیش بین ξ دارای رابطه زیر باشند

$$\begin{cases} \eta_1 = \gamma_{11}\xi + \zeta_1 \\ \eta_2 = \gamma_{21}\xi + \beta\eta_1 + \zeta_2 \end{cases} \quad (5.1)$$

به مدل (۵.۱) مدل ساختاری می گویند. صورت ماتریسی مدل (۵.۱) عبارتست از

$$\begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ \beta_{21} & 0 \end{pmatrix} \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} + \begin{pmatrix} \gamma_{11} \\ \gamma_{21} \end{pmatrix} (\xi) + \begin{pmatrix} \zeta_1 \\ \zeta_2 \end{pmatrix}$$

که می توان به صورت ساده تر زیر نوشت

$$\eta = B\eta + \Gamma\xi + \zeta \quad (6.1)$$

منظور از مدل ساختاری، صرفاً روابط علی بین متغیرهای پنهان است، هدف مدل ساختاری کشف اثرات مستقیم و غیرمستقیم متغیرهای پنهان برونزا بر متغیرهای پنهان درونزا می باشد. مدل معادلات ساختاری از کنار هم قرار دادن مدل های اندازه گیری و ساختاری تشکیل می شود. به عبارت دقیق تر صورت کلی مدل معادلات ساختاری عبارتست از

$$\begin{cases} X = \Lambda_x \xi + \delta \\ Y = \Lambda_y \eta + \varepsilon \\ \eta = B\eta + \Gamma\xi + \zeta \end{cases} \quad (7.1)$$

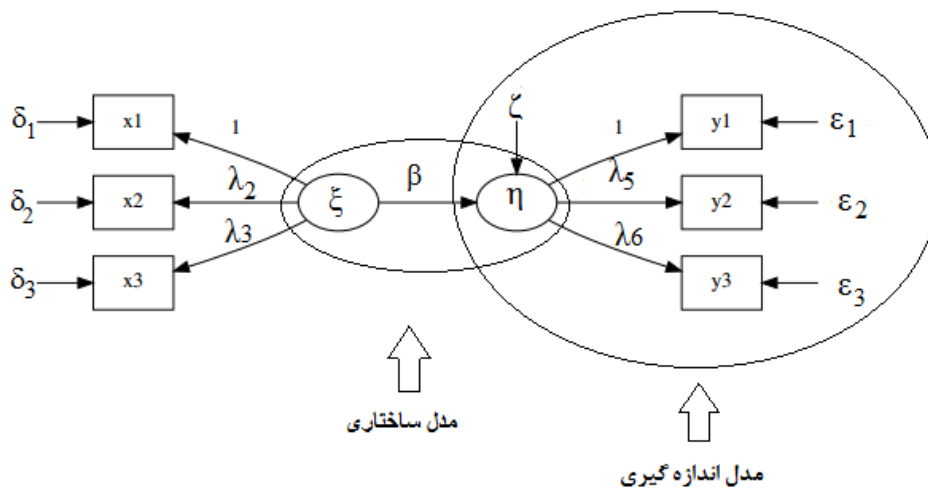
همان طور که قبلا دیدیم برای برازش یک مدل تحلیل عاملی باید ساختار واریانس-کواریانس مدل را برآورد کرد. در این جا نیز، برای برازش یک مدل معادلات ساختاری باید ابتدا ساختار واریانس-کواریانس مدل را تعیین کرد و سپس به برآورد پارامترهای مدل پرداخت. برای تعیین ماتریس واریانس-کواریانس مدل یک سری فرضیات را به صورت زیر در نظر می گیریم

$$\begin{cases} E(X) = E(Y) = 0 & , & E(\delta) = E(\varepsilon) = 0 \\ E(\eta) = E(\xi) = 0 & , & E(\xi\xi^T) = \Phi & , & E(\eta\eta^T) = \psi \\ E(\zeta) = 0 & , & E(\zeta\zeta^T) = \Psi & , & Cov(\varepsilon, \eta) = 0 \end{cases}$$

با توجه به فرضیات فوق، ساختار ماتریس واریانس-کواریانس مدل معادلات ساختاری (۷.۱) عبارتست از

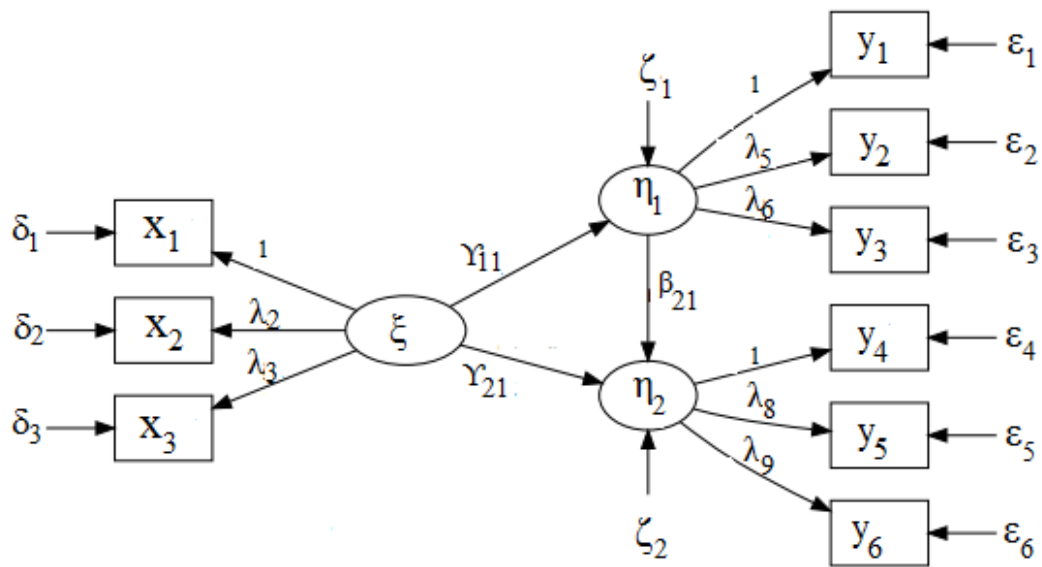
$$\Sigma(\theta) = \begin{pmatrix} \Lambda_x \Phi \Lambda_x^T & \Lambda_x \Phi \Gamma^T (I - B^T)^{-1} \Lambda_y^T \\ \Lambda_y (I - B)^{-1} \Gamma \Phi \Lambda_x^T & \Lambda_y (I - B)^{-1} (\Gamma \Phi \Gamma^T + \psi) (I - B^T)^{-1} \Lambda_y^T + \Theta_\varepsilon \end{pmatrix}$$

فرم نمودار مسیر مدل ساختاری و مدل اندازه گیری به صورت زیر است.



شکل ۱۰.۱: نمودار مسیر مدل معادلات ساختاری

شکل ۱۰.۱ یک نمونه از نمودار مسیر مدل معادلات ساختاری را نشان می دهد.



شکل ۱۱.۱: نمودار مدل معادلات ساختاری

و شکل ۱۱.۱ یک نمونه دیگر از نمودار مسیر مدل معادلات ساختاری را نشان می‌دهد. که رابطه‌های موجود در آن از مدل ریاضی زیر پیروی می‌کنند.

$$\mathbf{y} = \mathbf{B}\mathbf{y} + \mathbf{\Gamma}\mathbf{x} + \boldsymbol{\zeta} \quad (۸.۱)$$

و مدل اندازه‌گیری و مدل ساختاری آن به صورت زیر نوشته می‌شود

$$\begin{aligned} x_1 &= 1\xi + \delta_1 \\ x_2 &= \lambda_2\xi + \delta_2 \\ x_3 &= \lambda_3\xi + \delta_3 \end{aligned}$$

$$\begin{aligned} y_1 &= 1\eta_1 + \epsilon_1 \\ y_2 &= \lambda_5\eta_1 + \epsilon_2 \\ y_3 &= \lambda_6\eta_1 + \epsilon_3 \\ y_4 &= 1\eta_2 + \epsilon_4 \\ y_5 &= \lambda_8\eta_2 + \epsilon_5 \\ y_6 &= \lambda_9\eta_2 + \epsilon_6 \end{aligned}$$

$$X = \Lambda_x \xi + \delta$$

$$Y = \Lambda_y \eta + \varepsilon$$

فرم ماتریسی مدل اندازه گیری بالا به صورت زیر می باشد

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} (\xi) + \begin{pmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \end{pmatrix}$$

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \end{pmatrix} = \begin{pmatrix} 1 & \circ \\ \lambda_5 & \circ \\ \lambda_6 & \circ \\ \circ & 1 \\ \circ & \lambda_8 \\ \circ & \lambda_9 \end{pmatrix} \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \end{pmatrix}$$

که مدل ساختاری آن نیز به صورت زیر نوشته می شود

$$\eta_1 = \gamma_{11} \xi + \zeta_1$$

$$\eta_2 = \gamma_{21} \xi + \beta_{21} \eta_1 + \zeta_2$$

فرم ماتریسی مدل ساختاری آن نیز به صورت زیر نوشته می شود

$$\begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} = \begin{pmatrix} \circ & \circ \\ \beta_{21} & \circ \end{pmatrix} \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} + \begin{pmatrix} \gamma_{11} \\ \gamma_{21} \end{pmatrix} (\xi) + \begin{pmatrix} \zeta_1 \\ \zeta_2 \end{pmatrix}$$

$$\eta = B\eta + \Gamma\xi + \zeta$$

فرم ماتریسی مدل معادلات ساختاری شکل بالا به صورت زیر نوشته می شود.

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ \eta_1 \\ \eta_2 \\ \xi \end{pmatrix} = \begin{pmatrix} \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & 1 & \circ & \circ \\ \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \lambda_2 & \circ & \circ \\ \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \lambda_3 & \circ & \circ \\ \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & 1 & \circ \\ \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \lambda_5 & \circ \\ \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \lambda_6 & \circ \\ \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & 1 \\ \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \lambda_8 \\ \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \lambda_9 & \circ \\ \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \gamma_{11} \\ \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \beta_{21} & \gamma_{21} & \circ \\ \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ \eta_1 \\ \eta_2 \\ \xi \end{pmatrix} + \begin{pmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \\ \zeta_1 \\ \zeta_2 \\ \xi \end{pmatrix}$$

۳.۶.۱ مراحل مدل سازی معادلات ساختاری

همانند مدل رگرسیون چندگانه، مدل سازی معادلات ساختاری شامل چند مرحله است که به ترتیب به آن ها اشاره می کنیم.

تدوین مدل

قبل از هر نوع جمع آوری داده و تحلیل، پژوهشگر بایستی مدلی را تدوین نماید که مقادیر واریانس-کواریانس آن را تأیید نمایند. یک مدل هنگامی به خوبی تدوین شده است که مدل واقعی جامعه با مدل نظری فرض شده، سازگار باشد. به عبارت دیگر ماتریس کواریانس نمونه‌ای (S) باید به طور بسنده‌ای به وسیله مدل نظری تحت آزمون بازتولید شود. بنابراین هدف تحقق مدلی است که نزدیک‌ترین برازش را با ساختار کواریانس مدل دارا باشد. مثال ساده‌ای را با دو متغیر X و Y در نظر بگیرید. ما براساس پژوهش قبلی می دانیم که این دو متغیر با یکدیگر ارتباط دارند، اما چرا؟ و کدام ارتباط نظری بیانگر این رابطه است؟ آیا X بر Y اثر می گذارد یا عکس این حالت برقرار می باشد و یا متغیر سومی به نام Z بر هر دوی آن ها اثر می گذارد؟ گاه ممکن است با در نظر گرفتن مدل اولیه نامناسب، باعث حذف یک پارامتر مهم یا یک متغیر بااهمیت از مدل شویم. علاوه بر این ممکن است یک پارامتر یا متغیر نامناسب در مدل وارد شود که سبب ایجاد آریبی در برآورد پارامترها شده و نوعی خطا را در تدوین مدل به وجود آورد.

تشخیص مدل

در مدل‌سازی معادلات ساختاری حل مسئله تشخیص مدل پیش از برآورد پارامترها بسیار با اهمیت است. در تشخیص مدل این سؤال مطرح می‌شود که: آیا براساس داده‌های نمونه‌ای موجود در ماتریس کواریانس نمونه‌ای (S) و مدل نظری تعریف شده به‌وسیله ماتریس کواریانس جامعه (Σ) می‌توان مجموعه‌ی منحصر به فردی از برآورد پارامترها یافت؟

پیش از توضیح در مورد تشخیص مدل توضیحاتی را در مورد پارامترهای مدل ارائه می‌دهیم. هر پارامتر در مدل به سه صورت پارامتر آزاد، ثابت یا مقید مشخص می‌شود. پارامتر آزاد پارامتری است که شناخته شده نیست و نیازمند برآورد است. پارامتر ثابت، پارامتری است که آزاد نیست اما برای آن یک مقدار مشخص (به طور معمول مقدار صفر یا ۱) تعریف شده است. یک پارامتر مقید نیز پارامتری است که مشخص نیست اما برابر با یک یا تعداد بیشتری پارامتر است. یک پارامتر مقید به‌طور کامل آزاد نیست، اما تا حدودی یا به‌طور کامل محدود است. شناسایی در مدل‌سازی معادلات ساختاری به دو صورت، شناسایی بارهای واحد^{۵۶} (ULI) و شناسایی واریانس واحد^{۵۷} (UVI) انجام می‌گیرد که در (ULI) اغلب بار عاملی اولین شاخص مقدار ثابت یک، می‌باشد که به آن شاخص مرجع می‌گویند و

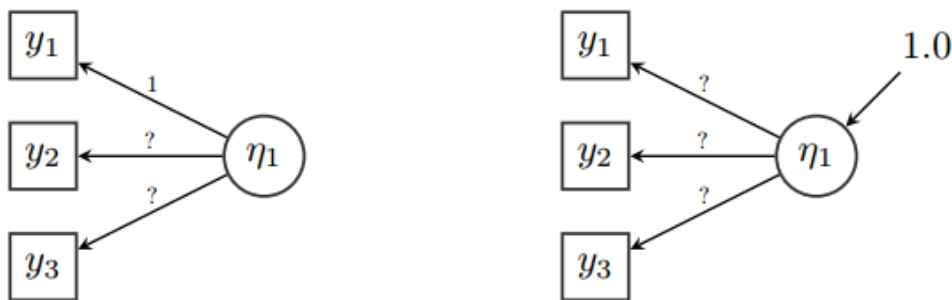
$$\Lambda = \begin{pmatrix} 1 & \circ \\ \lambda_{21} & \circ \\ \lambda_{31} & \circ \\ \circ & 1 \\ \circ & \lambda_{52} \\ \circ & \lambda_{62} \end{pmatrix}, \quad \Psi = \begin{pmatrix} \psi_{11} & - \\ \psi_{21} & \psi_{22} \end{pmatrix} \quad (9.1)$$

در (UVI) واریانس عامل برابر با مقدار ثابت یک می‌باشد.

$$\Lambda = \begin{pmatrix} \lambda_{11} & \circ \\ \lambda_{21} & \circ \\ \lambda_{31} & \circ \\ \circ & \lambda_{42} \\ \circ & \lambda_{52} \\ \circ & \lambda_{62} \end{pmatrix}, \quad \Psi = \begin{pmatrix} 1 & - \\ \psi_{21} & 1 \end{pmatrix} \quad (10.1)$$

^{۵۶} Unit loading identification

^{۵۷} Unit variance identification



شکل ۱۲.۱: نمودار مدل شناسایی در مدل سازی معادلات ساختاری

شکل ۱۲.۱ یک نمونه از مدل شناسایی در مدل سازی معادلات ساختاری را نشان می دهد. در بسیاری از مدل ها مهم نیست که از کدام روش استفاده کنیم؛ در تحلیل گروه های چندگانه ی مدل سازی معادلات ساختاری (SEM) معمولاً از شناسایی بارهای واحد (ULI) استفاده می کنیم.

تشخیص مدل در واقع به طرح پارامترها به عنوان ثابت، آزاد یا مقید بستگی دارد. پس از آن که مدل و پارامترها تدوین شدند، این پارامترها برای شکل دادن به یک و تنها یک Σ با یکدیگر ترکیب می شوند. اگر دو یا تعداد بیشتری از مجموعه پارامترها ماتریس Σ یکسانی را تولید کنند، آن گاه این مجموعه ها معادل یا همتا خوانده می شوند.

سه سطح برای تشخیص مدل وجود دارد که در زیر به بیان آن ها می پردازیم. مدلی فروم مشخص^{۵۸} است که، یک یا تعداد بیشتری از متغیرها نتوانند به طور یکتایی مشخص شوند، زیرا اطلاعات کافی در ماتریس S وجود ندارد. مدلی کاملاً مشخص است که، همه پارامترها به دلیل وجود اطلاعات کافی در ماتریس S به طور منحصر به فردی تعیین شوند. مدلی فرامشخص^{۵۹} است که، در آن بیش از یک جواب برای یک یا چند پارامتر وجود دارد. اگر مدل فروم مشخص باشد برآورد پارامترها قابل اعتماد نبوده و در چنین حالتی درجات آزادی مدل صفر یا منفی است. این مدل ممکن است با افزودن قیدهایی مشخص شود. مدل های کاملاً مشخص و فرامشخص برای برآورد پارامترها مناسب هستند.

برآورد مدل

گام بعدی به دست آوردن برآوردهایی برای هریک از پارامترهای تعیین شده در مدل است. به عبارتی باید ماتریس $\Sigma = \Sigma(\theta)$ را برآورد کنیم. برآورد پارامترها باید به گونه ای باشد که نزدیک ترین ماتریس به ماتریس واریانس-کواریانس نمونه ای بازتولید شود و خطا یعنی $S - \Sigma$ کمینه شود. برخی از روش های اولیه برای این منظور شامل حداقل مربعات غیر وزنی^{۶۰}،

^{۵۸} Under identified

^{۵۹} Over-identified

^{۶۰} Unweight least square

حداقل مربعات معمول^{۶۱}، حداقل مربعات تعمیم یافته^{۶۲} و روش حداکثر درستنمایی^{۶۳} است. از میان این روش ها تنها روش حداقل مربعات غیروزی وابسته به مقیاس است. در حقیقت به دنبال آن مقداری از θ هستیم که تفاوت بین آنچه که ما در داده ها مشاهده می کنیم (S) و آنچه که در مدل پیداست $\Sigma(\theta)$ کمینه کند. سپس مقادیر برآورد شده به وسیله $\hat{\theta}$ نشان داده می شوند. و برآورد ماتریس کواریانس مدل ضمنی می تواند به صورت $\hat{\Sigma} = \Sigma(\hat{\theta})$ نوشته شود. راه های بسیاری برای تعیین این تفاوت وجود دارد. مشابه روش تحلیل عاملی در این مدل نیز از تابع اختلاف زیر استفاده می کنیم.

$$F_{ML}(\theta) = tr [R\Sigma^{-1}(\theta)] - \log|R\Sigma^{-1}(\theta)| - p \quad (11.1)$$

در این صورت $\hat{\Sigma} = \Sigma(\hat{\theta}) = \underset{\theta}{argmin} F_{ML}(\theta)$ برقرار می باشد.

آزمون مدل

پس از آن که برآورد پارامترها برای یک مدل تدوین شده و مشخص به دست آمد، محقق باید تعیین کند که داده ها تا چه حد با مدل برازش دارند؟

دو شیوه برای بررسی برازش مدل وجود دارد، ابتدا ملاحظه برخی آزمون های عمومیت یافته برای برازش کل مدل است و شیوه دوم بررسی برازش پارامترهای منفرد (تکی) در هر یک از اجزای مدل است. آزمون های کلی با عنوان معیارهای برازش مدل شناخته می شوند. بسیاری از این شاخص ها بر مبنای مقایسه ماتریس کواریانس اقتباس شده از مدل Σ با ماتریس کواریانس نمونه ای S ساخته شده اند.

برای بررسی برازش پارامترهای منفرد سه آزمون اصلی مورد استفاده قرار می گیرند. اول آن که آیا یک پارامتر آزاد به طور معناداری با صفر تفاوت دارد یا خیر؟ دوم آن که آیا علامت پارامتر با آنچه به لحاظ نظری مورد انتظار بوده هماهنگ است؟ و سوم اینکه برآورد پارامترها باید در دامنه مقادیر مورد انتظار قرار گیرند. هر یک از این سؤالات با کمک روش ها و آزمون های آماری مناسب پاسخ داده می شوند.

۷.۱ شاخص های برازش

مدل تدوین شده بر مبنای چارچوب نظری و پیشینه تجربی تا چه اندازه با واقعیت انطباق دارد؟ این سؤالی است که هر پژوهشگری که پژوهش خود را در قالب یک مدل معادله ساختاری مدون کرده است با آن روبه رو است و مایل است به آن پاسخ دهد. برای پاسخ به این سوال از شاخص های نیکویی برازش مدل استفاده می شود. سه گروه کلی برازش مدل عبارتند از

^{۶۱} Usual least square

^{۶۲} Generalized least square

^{۶۳} Maximum likelihood

۱- شاخص برازش مطلق ۲- شاخص برازش تطبیقی ۳- شاخص برازش مقتصد. در ادامه توضیح مختصری در خصوص این شاخص‌ها ارائه می‌گردد که برگرفته از قاسمی (۱۳۹۲) است.

۱.۷.۱ شاخص‌های برازش مطلق

شاخص‌های برازش مطلق^{۶۴} شاخص‌هایی هستند که بر مبنای تفاوت واریانس‌ها و کواریانس‌های مشاهده شده از یک طرف و واریانس‌ها و کواریانس‌های پیش‌بینی شده بر مبنای پارامترهای مدل تدوین شده از طرف دیگر قرار دارند.

معیار برازش در این گروه از شاخص‌ها، مقایسه مدل با مدل‌های رقیب نیست (هر آنچه که در شاخص‌های تطبیقی مورد نظر است) و وابسته به تعداد پارامترهایی که در مدل توسط پژوهشگر تعریف شده‌اند (هر آنچه که در شاخص‌های مقتصد مورد نظر است) نمی‌باشد.

با توجه به این که افزودن هر پارامتر به مدل (تا جایی که برآورد پارامتر به لحاظ محاسباتی امکان‌پذیر باشد) باعث نزدیک‌تر شدن دو گروه واریانس‌ها و کواریانس‌های ذکر شده می‌شود، اضافه کردن هر پارامتر به مدل باعث بهبود شاخص‌های برازش مطلق می‌شود تا جایی که اگر پژوهشگر حداکثر پارامترهای ممکن را تعریف کند (مدل اشباع^{۶۵} شود) پژوهشگر به برازش کامل دست خواهد یافت که در این جا به معنای برابر شدن واریانس‌ها و کواریانس‌های مشاهده شده و باز تولید شده به وسیله مدل می‌باشد. برخی از این شاخص‌ها به شرح زیر می‌باشند.

کای-دو

شاخص کای-دو (χ^2) را می‌توان به عنوان عمومی‌ترین و پرکاربردترین شاخص برازش در مدل‌سازی معادله ساختاری تلقی کرد. کای-دو مدل به نام‌های کای-دو نسبت درست‌نمایی^{۶۶}، شاخص برازش کای-دو^{۶۷}، شاخص بدی برازش کای-دو^{۶۸}، خوانده می‌شود. علاوه بر کای-دو مدل، مقادیر دیگری که اصلاح شده‌ی آن هستند در موقعیت‌های خاصی محاسبه می‌شوند که از جمله مهم‌ترین آن‌ها می‌توان به مقادیر کای-دو نسبی^{۶۹} (کای-دو بهنجار^{۷۰} یا کای-دو بهنجار شده^{۷۱}) به عنوان یکی از شاخص‌های مقتصد نام برد. کای-دو تصحیح شده ساتورا-

^{۶۴} Absolute fit indices

^{۶۵} مدلی که در آن همه‌ی مسیرهای ممکن ترسیم شده‌اند، می‌توان آن را نقطه مقابل مدل استقلال دانست.

^{۶۶} Likelihood ratio chi-square

^{۶۷} Fitness index of chi-square

^{۶۸} Baddness-of-fit index of chi-square

^{۶۹} Relative chi-square

^{۷۰} Normal chi-square

^{۷۱} Normed chi-square

بنتلر^{۷۲} نیز که در موقعیت‌های خاص به‌ویژه عدم برقراری نرمال بودن چندمتغیره قابل استفاده است در واقع اصلاح شده‌ی کای-دو مدل است.

اولین نکته‌ای که درباره تفسیر مقدار کای-دو می‌توان اظهار کرد این است که هرچه مقدار آن کوچک‌تر باشد برازش داده‌ها به مدل بهتر است تا جایی که مقدار صفر برای آن نشانه‌ی برازش کامل است. مقدار کای-دو به لحاظ نظری در دامنه‌ی بین صفر تا بی‌نهایت تغییر می‌کند. هم‌چنین می‌توان گفت در مدل‌سازی و در یک موقعیت پژوهشی خاص مقدار کای-دو بین کای-دو مدل اشباع شده تا کای-دو مدل استقلال تغییر می‌کند.

مبنای محاسبه مقدار کای-دو تفاوتی است که بین دو ماتریس وجود دارد: اول ماتریس واریانس-کواریانس نمونه‌ای که حاصل محاسبه واریانس متغیرهای مشاهده شده‌ی حاضر در مدل و کواریانس بین کلیه زوج‌های ممکن بین این متغیرها است و برآوردی از واریانس و کواریانس‌های متغیرهای مورد پژوهش در جامعه آماری تلقی می‌شود. دوم ماتریس واریانس-کواریانس باز تولید شده بر مبنای پارامترهای برآورد شده در مدل تدوین شده یا مفروض که در واقع برآوردی از ساختار واریانس-کواریانس متغیرهای مورد مطالعه بر روی یک جامعه آماری نامحدود است. می‌توان این مبنا را در یک فرآیند سه مرحله‌ای تشریح کرد. در مرحله‌ی اول ماتریس مربعی با مرتبه n براساس داده‌های نمونه‌ای گردآوری شده تشکیل می‌شود. مرتبه ماتریس منطبق با تعداد متغیرهای مشاهده شده حاضر در مدل تدوین شده است. به این ترتیب اگر چهار متغیر مشاهده شده در مدل وجود دارد با یک ماتریس مربع^{۷۳} مرتبه ۴ روبه‌رو هستیم که عناصر قطری آن را واریانس‌های چهار متغیر مشاهده شده و عناصر غیرقطری آن را کواریانس‌های بین این متغیرها به‌صورت دوجه‌دو تشکیل می‌دهند. ماتریس واریانس-کواریانس نمونه‌ای که اغلب با نماد S نمایش داده می‌شود برای چهار متغیر مشاهده شده به‌شرح زیر است

$$S = \begin{bmatrix} s_{11} & s_{12} & s_{13} & s_{14} \\ s_{21} & s_{22} & s_{23} & s_{24} \\ s_{31} & s_{32} & s_{33} & s_{34} \\ s_{41} & s_{42} & s_{43} & s_{44} \end{bmatrix} \quad (12.1)$$

در مرحله‌ی دوم بر مبنای داده‌های موجود در ماتریس کواریانس نمونه‌ای، پارامترهای آزادی که در مدل تدوین شده وجود دارند برآورد می‌شوند. روش‌های برآورد مختلفی وجود دارد که هر کدام دارای کاربرد خاص خود هستند عمومی‌ترین آن‌ها روش ماکزیمم درست‌نمایی (ML) است. مهم‌ترین پیش فرض‌های این روش کمی و پیوسته بودن متغیرهای حاضر در مدل و برقراری نرمال بودن چندمتغیره است. در مرحله سوم با استفاده از تکنیک ترکیب‌زدایی

^{۷۲} Satoora-Bentler corrected chi-square

^{۷۳} منظور ماتریسی است که در آن $m = n$ یا به‌طور ساده تعداد سطرها و ستون‌ها برابر است.

کواریانس‌ها^{۷۴} (مولر^{۷۵}، ۲۰۰۵) ماتریس واریانس-کواریانس بر مبنای مقادیر برآورد شده برای پارامترها بازسازی می‌شود. این مقادیر در واقع برآوردی از این ماتریس در جامعه آماری نامحدود است. در این مرحله هر چه پارامترهای تعریف شده به‌عنوان پارامتر ثابت، آزاد یا مقید قابل قبول‌تر باشند، ماتریس کواریانس باز تولید شده به ماتریس کواریانس نمونه‌ای یا مشاهده شده نزدیک‌تر خواهد بود. به عبارتی حاصل تفریق آن‌ها یا ماتریس باقی‌مانده به یک ماتریس صفر^{۷۶} نزدیک‌تر می‌شود. ماتریس جامعه نامحدود که با استفاده از مقادیر پارامترهای مدل تدوین شده برآورد می‌شود را می‌توان به شکل زیر نمایش داد

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} & \sigma_{24} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 & \sigma_{34} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_4^2 \end{bmatrix} \quad (13.1)$$

هر چه مقادیر قرار گرفته در سلول‌های دو ماتریس S و Σ به یکدیگر نزدیک‌تر باشند و مقادیر موجود در عناصر ماتریس باقیمانده ناشی از تفاضل دو ماتریس به صفر نزدیک‌تر شوند مقادیر کای-دو نیز کاهش خواهد یافت تا جایی که اگر این دو ماتریس معادل یا برابر^{۷۷} باشند و بنابراین ماتریس باقیمانده یک ماتریس صفر باشد مقدار کای-دو نیز برابر با صفر خواهد شد.

$$S - \Sigma = \circ \Rightarrow \chi^2 = \circ, Sig = 1$$

با توجه به توضیحات مقدماتی که داده شده امکان تشریح فرمول محاسبه‌ی کای-دو تا حد زیادی فراهم می‌شود.

$$\chi^2 = (n - 1) \times F_{ML}$$

که در آن F_{ML} به صورت زیر تعریف می‌شود

$$F_{ML} = \log(|\Sigma|) + tr(S\Sigma^{-1}) - \log(|S|) - p \quad (14.1)$$

که در آن $\log(|S|)$ نمادی برای لگاریتم طبیعی دترمینان ماتریس کواریانس مشاهده شده و $\log(|\Sigma|)$ نمادی برای لگاریتم طبیعی دترمینان ماتریس کواریانس باز تولید شده است. $tr(S\Sigma^{-1})$ نیز اثر ماتریس^{۷۸} حاصل از ضرب دو ماتریس کواریانس مشاهده شده و وارون ماتریس کواریانس باز تولید شده است.

^{۷۴}Decomposition of covariance

^{۷۵}Muller

^{۷۶}Zero matrix

^{۷۷}Equal matrices

^{۷۸}اثر ماتریس حاصل جمع مقادیر قرار گرفته در قطر اصلی ماتریس است.

ریشه دوم میانگین مربعات باقیمانده

ماتریس باقیمانده هم می‌تواند برای ارزیابی برازش کلی (مدل تدوین شده) و هم برازش جزئی (پارامتر تعریف شده بین دو متغیر) مورد استفاده قرار گیرد. این شاخص یکی از شاخص‌های برازش است که از آن برای مقایسه دو مدل متفاوت با داده‌های یکسان استفاده می‌شود. حداقل مقدار برای این شاخص صفر است یعنی حالتی که برازش کامل و ماتریس باقیمانده یک ماتریس صفر (ماتریسی که همه‌ی درایه‌های آن صفر هستند) می‌باشد. اما حداکثر آن با توجه به مقدار کواریانس‌ها در ماتریس باقیمانده می‌تواند عددی کوچک یا بزرگ باشد. هر چه مقدار ریشه دوم میانگین مربعات باقیمانده (RMR) کوچک‌تر باشد برای یک مدل می‌تواند معیاری برای برتری آن نسبت به مدل‌های دیگر باشد. این شاخص از طریق فرمول زیر محاسبه می‌شود

$$RMR = \sqrt{\frac{\sum_{ij} (S_{ij} - \sigma_{ij})^2}{k}}$$

که در آن S_{ij} شامل کلیه عناصر غیرزاید در ماتریس کواریانس مشاهده شده و σ_{ij} شامل کلیه عناصر غیرزاید در ماتریس کواریانس بازسازی شده است. اگر ماتریس باقیمانده را با D نمایش دهیم آن‌گاه می‌توان به جای $S_{ij} - \sigma_{ij}$ به اختصار از D_{ij} استفاده کرد که نشان‌دهنده‌ی کلیه عناصر غیرزاید در ماتریس باقیمانده است. k نیز تعداد کل عناصر غیرزاید در ماتریس کواریانس مشاهده شده است.

دیوید ای. کنی تاکید می‌کند که شاخص نیکویی برازش (GFI) و شاخص نیکویی برازش (AGFI) به شدت تحت تاثیر حجم نمونه هستند و برای مدل‌های تدوین شده‌ی بسیار ضعیف می‌توانند مقادیر بالایی را نشان دهند. به دلیل ضعف این شاخص‌ها از توضیح درباره‌ی آن‌ها در این قسمت خودداری می‌کنیم.

۲.۷.۱ شاخص‌های برازش تطبیقی

شاخص برازش تطبیقی^{۷۹} در واقع گامی در جهت تکمیل شاخص‌های برازش مطلق محسوب می‌شوند به این ترتیب که با مبنا قرار دادن یک یا چند مدل مدل نظری تدوین شده تحت آزمون را با آن مقایسه و نشان می‌دهد که آیا به لحاظ آماری قابل قبول‌تر تلقی می‌شود، ضعیف‌تر است و یا این که تفاوتی با آن ندارد. این مدل مبنا به‌طور پیش فرض یک مدل استقلال^{۸۰} است که در آن تنها پارامتر آزاد، واریانس متغیرهای مشاهده شده است. بنابراین در مجموع این شاخص‌ها نشان‌دهنده‌ی آن هستند که مدل تدوین شده تا چه اندازه توانسته است از یک مدل استقلال

^{۷۹}Comparative fit indices

^{۸۰}مدلی که در آن متغیرهای موجود فاقد هرگونه رابطه یک‌سویه یا دوسویه با یکدیگرند و کواریانس میان آن‌ها صفر در نظر گرفته شده است، که به آن‌ها مدل صفر می‌گویند و مبنایی برای ارزیابی مدل مفروض را فراهم می‌آورند.

به عنوان مدلی که متغیرهای آن فاقد هرگونه کواریانس یا همبستگی با یکدیگر هستند دور شود (فاصله بگیرد).

در بیشتر موارد مدل مبنا یک مدل مستقل است که تنها پارامترهای آزاد آن واریانس‌های متغیرهای مشاهده است و هیچ پارامتر دیگری نظیر کواریانس میان متغیرها، بارهای عاملی در مدل‌های اندازه‌گیری و یا اثر متغیرهای بیرونی بر درونی و یا درونی بر درونی در آن تعریف نشده است. شاخص‌های برازش تطبیقی نشان می‌دهند که مدل تدوین شده تا چه اندازه توانسته از یک مدل مستقل فاصله بگیرد. هرچه این فاصله بیشتر باشد برازش مدل مطلوب‌تر تلقی می‌شود. مدل استقلال یکی از انواع مدل‌های صفر^{۸۱} است.

باید توجه داشت که در محاسبه‌ی این شاخص‌ها از مقادیر کای-دو مدل تدوین شده و کای-دو مدل استقلال استفاده می‌شود. برای این دو مقدار به ترتیب از نمادهای χ^2_n و χ^2_m استفاده می‌شود. به همین ترتیب df_n و df_m به ترتیب نمادهایی برای درجه آزادی مدل تدوین شده و درجه آزادی مدل استقلال است. مهمترین شاخص‌های برازش تطبیقی به شرح زیر هستند.

شاخص برازش هنجار شده بنتلر-بونت

این شاخص اولین بار توسط بنتلر و بونت^{۸۲} (۱۹۸۰) در مقاله‌ای با عنوان آزمون‌های معناداری و نیکویی برازش در تحلیل ساختارهای کواریانس طرح شد. اصلی‌ترین نقطه ضعف این شاخص، عدم حساسیت آن به افزودن پارامتر به مدل است به نحوی که هرچه پارامتر به مدل افزوده شود مقدار این شاخص نیز افزایش می‌یابد (به نقل از قاسمی، ۱۳۹۲). به همین دلیل کمتر از این شاخص استفاده می‌شود. شاخص برازش هنجار شده بنتلر-بونت (NFI) یا ($\Delta 1$) از طریق فرمول زیر محاسبه می‌شود.

$$NFI = \frac{\chi^2_n - \chi^2_m}{\chi^2_n}$$

مقدار قابل قبول برای این شاخص حداقل 0.90 و مقداری که نشان‌دهنده‌ی یک برازش خوب است حداقل 0.95 در نظر گرفته می‌شود. اگر مقداری کمتر را اختیار کند مدل ما نیاز به اصلاح دارد.

شاخص برازش نسبی

شاخص برازش نسبی (RFI) یا (ρ_1) توسط مارش^{۸۳} و همکاران (۱۹۸۸) طرح شده است. مقدار این شاخص بین صفر تا ۱ تغییر می‌کند به طوری که مقادیر نزدیکتر به ۱ به عنوان برازش بهتر

^{۸۱}Null models

^{۸۲}Bentler and Bonnet

^{۸۳}Marsh, H. W.

داده‌ها در مدل تفسیر می‌شود. مقدار آن از طریق فرمول زیر محاسبه می‌شود.

$$RFI = 1 - \frac{\chi_m^2/df_m}{\chi_n^2/df_n}$$

شاخص برازش افزایشی

شاخص برازش افزایشی (IFI) یا $(\Delta\chi^2)$ اولین بار توسط بولن^{۸۴} (۱۹۸۹) طرح شده است. مقدار این شاخص از طریق فرمول زیر محاسبه می‌شود.

$$IFI = \frac{\chi_n^2/\chi_m^2}{\chi_n^2/df_m} \quad (15.1)$$

شاخص توکر-لویس

شاخص توکر-لویس (TLI) یا $(\rho\chi^2)$ تلاش می‌کند تا نقطه ضعف شاخص بنتلر-بونت را در به حساب نیابردن جریمه شاخص برای افزودن پارامتر را مرتفع سازد. این شاخص بر مبنای متوسط ضرایب همبستگی بین متغیرها در مدل قرار دارد. هرچه این ضرایب کوچکتر باشند شاخص توکر-لویس نیز مقدار کوچکتری را نشان خواهد داد.

از این شاخص در دو جهت می‌توان استفاده کرد: اول مقایسه‌ی دو یا چند مدل متفاوت با داده‌های یکسان و دوم مقایسه‌ی مدل مفروض یا یک مدل صفر نظیر مدل استقلال. مقدار این شاخص بین صفر تا ۱ تغییر می‌کند و مقدار ۰/۹۵ یا بیشتر منعکس کننده‌ی یک مدل خوب است. اگر این شاخص بزرگتر از ۱ شود برابر با ۱ قرار داده خواهد شد. این شاخص به نام شاخص برازش هنجار نشده یا NNFI نیز خوانده می‌شود. مقدار این شاخص از طریق فرمول زیر محاسبه می‌شود.

$$TLI = \frac{(\chi_n^2/df_n) - \chi_m^2/df_m}{(\chi_n^2/df_n) - 1}$$

شاخص برازش تطبیقی

شاخص برازش تطبیقی (CFI) نیز بر مبنای همبستگی بین متغیرهای حاضر در مدل قرار دارد، به نحوی که ضرایب بالای همبستگی بین آن‌ها به مقادیر بالای شاخص برازش تطبیقی می‌انجامد. مقدار این شاخص بین صفر تا ۱ تغییر می‌کند و مقدار ۰/۹۵ یا بیشتر منعکس کننده‌ی یک مدل خوب است. اگر این شاخص بزرگتر از ۱ شود برابر با ۱ قرار داده خواهد شد. مقدار این شاخص از طریق فرمول زیر محاسبه می‌شود.

$$CFI = \frac{(\chi_n^2 - df_n) - (\chi_m^2 - df_m)}{(\chi_n^2 - df_n)}$$

^{۸۴}Bollen

۳.۷.۱ شاخص‌های برازش مقتصد

بحث اقتصاد مدل به موضوع هزینه/پاداش مربوط می‌شود و طرح آن به این دلیل اساسی بوده که پژوهشگر در تدوین مدل خود، به‌منظور آزاد گذاشتن پارامترها اجتناب می‌کند و پارامترهایی را برای برآورد آزاد می‌گذارد که دارای مبانی نظری و تجربی قوی باشند. مبنای اصلی در طرح این شاخص‌ها تکیه آن‌ها بر درجه آزادی است. هر پارامتری که برای برآورد آزاد گذاشته می‌شود منجر به کاهش یک درجه آزادی برای مدل تدوین شده می‌شود. در نتیجه کاهش درجه آزادی مدل را می‌توان به‌عنوان هزینه‌ای قلمداد کرد که در اثر آزاد گذاشتن پارامترها به مدل تحمیل می‌شود. پاداش اخذ شده در مقابل هزینه تحمیل شده کاهش می‌دهد که در کای-دو مدل و در نتیجه بهبود شاخص‌های برازش مطلق و اغلب شاخص‌های برازش تطبیقی حاصل می‌شود. شاخص‌های برازش مقتصد^{۸۵} در واقع معیاری به‌دست محقق می‌دهند که آیا پاداش به‌دست آمده در مقایسه با هزینه از دست رفته توجیه دارد یا خیر.

با ارائه شاخص‌های برازش مقتصد تلاش شده تا اصلی‌ترین نقطه ضعف شاخص‌های برازش مطلق، یعنی بهبود مقدار شاخص‌های برازش با افزایش پارامتر به مدل جبران شود. مبنای اصلی در این گروه از شاخص‌های برازش آن است که به ازای هر پارامتری که به مدل اضافه می‌شود این شاخص‌ها جریمه می‌شوند.

شاخص نسبت اقتصاد

شاخص نسبت اقتصاد (PRATIO)، یک شاخص برازش محسوب نمی‌شود اما نشان می‌دهد که پژوهشگر تا چه حد در تعریف پارامترهای آزاد هزینه کرده است. این شاخص که بر مبنای نسبت درجه آزادی مدل تدوین شده به درجه آزادی مدل استقلال به‌دست می‌آید مقداری بین صفر تا ۱ به خود می‌گیرد. هرچه مقدار این شاخص کوچکتر باشد حاکی از هزینه‌ی بیشتری است که پژوهشگر در آزاد کردن پارامترها صرف کرده است. مقدار این شاخص از طریق فرمول زیر محاسبه می‌شود.

$$PRATIO = df_m / df_n$$

این نسبت در محاسبه‌ی سایر شاخص‌های مقتصد مورد استفاده قرار می‌گیرد.

شاخص برازش هنجار شده مقتصد

هرچه مدل تدوین شده به یک مدل اشباع شده نزدیک‌تر شود و در نتیجه درجه آزادی آن به صفر متمایل شود شاخص برازش هنجار شده جریمه می‌شود. به عبارتی شاخص برازش هنجار شده‌ی مقتصد (PNFI) از حاصل ضرب نسبت اقتصاد مدل در شاخص برازش هنجار شده حاصل

^{۸۵}Parsimonious fit indices

می شود.

$$PNFI = PRATIO \times NFI$$

$$PNFI = (df_m/df_n) \times NFI$$

مقادیر بالای ۰/۵ را برای این شاخص مناسب تلقی می کنند هرچند توافقی در این باره وجود ندارد. مقادیر بالاتر برای این شاخص حاکی از قابل قبول تر بودن مدل است با تاکید بر این نکته که مقدار قابل قبول برای آن ها به هیچ وجه قابل مقایسه با مقادیر قابل قبول برای شاخص های برازش تطبیقی (۰/۹۰ یا ۰/۹۵) نیست.

شاخص برازش تطبیقی مقتصد

می توان شاخص برازش تطبیقی مقتصد (PCFI) را شاخص برازش تطبیقی دانست که هرچه مدل به یک مدل اشباع شده نزدیک تر شود بیشتر جریمه می شود. مقادیر بالای ۰/۵ را برای این شاخص مناسب تلقی می کنند، هرچند توافقی در این باره وجود ندارد. فرمول آن به شرح زیر است.

$$PCFI = PRATIO \times CFI$$

$$PCFI = (df_m/df_n) \times CFI$$

ریشه دوم میانگین مربعات خطای برآورد

ریشه دوم میانگین مربعات خطای برآورد (RMSEA) برای اولین بار توسط اشتایگر^{۸۶} (۱۹۹۰) معرفی شده است. مانند RMR این شاخص نیز بر مبنای تحلیل ماتریس باقیمانده قرار دارد با این حال خصوصیت های مهم تری در مقایسه با آن نیز دارا می باشد.

بر خلاف بسیاری از شاخص های برازش دیگر در مدل سازی که تنها دارای برآورد نقطه ای هستند این شاخص برای فواصل اطمینان مختلف نیز قابل محاسبه است که می توان با استفاده از آن ها مشخص کرد که آیا مقدار بدست آمده برای مدل تدوین شده با مقدار ۰/۰۵ تفاوت معنادار دارد یا خیر؟ مقداری که اغلب به عنوان نقطه ی برش برای مدل های خوب و بد شناخته می شود. این شاخص که بر مبنای پارامتر غیر مرکزی قرار دارد بر اساس فرمول زیر محاسبه می شود.

$$RMSEA = \sqrt{\frac{\chi_m^2 - df_m}{(n - 1) \times df_m}}$$

که در آن n حجم نمونه و df_m درجه آزادی مدل تدوین شده است. اگر کای-دو مدل (χ_m^2) از درجه آزادی مدل کوچکتر باشد مقدار این شاخص برابر صفر در نظر گرفته می شود. مدل های قابل قبول دارای مقدار ۰/۰۵ یا کوچکتر برای این شاخص هستند. برازش مدل هایی که دارای

^{۸۶}Steiger, J. H.

مقادیر بالاتر از ۱/۰ هستند ضعیف برآورد می‌شود. توجه داشته باشید که مقدار این شاخص هنگامی که درجه آزادی کوچک است و حجم نمونه نیز بزرگ نیست می‌تواند گمراه کننده باشد. این شاخص در نرم‌افزارهای مختلف از جمله Amos اغلب با مقدار PCLOSE گزارش می‌شود. این مقدار فرضیه‌ی صفری را آزمون می‌کند که بر مبنای آن RMSEA برابر ۰/۰۵ است.

$$H_0 : RMSEA = 0/05$$

اگر مقدار PCLOSE بزرگتر از ۰/۰۵ باشد می‌توان نتیجه گرفت که مدل تدوین شده به یک مدل قابل قبول نزدیک است.

این شاخص به‌عنوان یکی از شاخص‌های مقتصد شناخته می‌شود که مقدار آن به دلیل پیچیدگی مدل اصلاح می‌شود. نتیجه می‌گیریم که در دو مدل مشخص با توان تبیین کنندگی یکسان برای داده‌های مشابه، مدل ساده‌تر ارجحیت دارد.

اگر مقدار این شاخص برابر صفر باشد تنها نشان می‌دهد که کای-دو مدل از درجه آزادی مدل کوچکتر است نه این که کای-دو لزوماً صفر باشد. نرم‌افزارهای کامپیوتری از جمله Amos فاصله اطمینان ۹۰٪ را برای این شاخص گزارش می‌کنند. چنانچه حد پایین این فاصله اطمینان کوچکتر از ۰/۰۵ باشد مدل تدوین شده به واقعیت جامعه آماری نزدیک تلقی می‌شود. در حالت آرمانی انتظار داریم که حد پایین این شاخص برابر صفر باشد. در هر حال وجود حجم نمونه اندک می‌تواند به حالتی پیچیده تبدیل شود که حد پایین شاخص از ۰/۰۵ کوچکتر و حد بالای آن از ۱/۰ بزرگتر باشد.

کای-دو بهنجار یا نسبی

یکی از شاخص‌های عمومی برای به حساب آوردن پارامترهای آزاد در محاسبه شاخص برازش، کای-دو بهنجار یا نسبی است که از تقسیم ساده‌ی مقدار کای-دو بر درجه آزادی مدل محاسبه می‌شود و با برجسب CMIN/DF گزارش می‌شود. اغلب مقادیر بین ۲ تا ۳ را برای این شاخص قابل قبول می‌دانند با این حال دیدگاه‌ها در این باره متفاوت است.

شوماخر و لومکس^{۸۷} (۱۳۸۸) مقادیر بین ۱ تا ۵، کارماینز و مک ایور^{۸۸} (۱۹۸۱) مقادیر بین ۲ تا ۳، اولمن^{۸۹} (۲۰۰۱) مقادیر بین ۱ تا ۲ و کلاین^{۹۰} (۲۰۰۵) مقادیر بین ۱ تا ۳ را قابل قبول می‌دانند.

^{۸۷} Schumacher and Lumex

^{۸۸} Carmenes and McEvire

^{۸۹} Olman

^{۹۰} Klein

۴.۷.۱ سایر شاخص‌ها

برخی از شاخص‌ها توسط برخی نویسندگان در گروه شاخص‌های مقتصد طبقه‌بندی شده‌اند چرا که مقدار آن‌ها به‌خاطر افزودن پارامتر جریمه می‌شوند. در هر حال آن‌ها دارای نقطه برش مشخصی نیستند و اغلب برای مقایسه مدل‌ها به‌کار می‌روند تا این که شاخصی برای ارزیابی یک مدل خاص باشند.

معیار اطلاع آکائیک و سایر شاخص‌های مرتبط

گروهی از شاخص‌های برازش مقتصد طرح شده‌اند که اغلب برای تحلیل یک مدل خاص به‌کار نمی‌روند بلکه از آن‌ها برای مقایسه‌ی مدل‌های مختلف استفاده می‌شوند. این شاخص‌ها عبارتند از

- معیار اطلاع آکائیک (AIC) اولین بار توسط ایچ. آکائیک^{۹۱} (۱۹۸۷) مطرح شد. فرمول این شاخص به‌شرح زیر می‌باشد

$$AIC = \chi_m^2 + 2q$$

- معیار براون-کادک^{۹۲} (BCC) نیز توسط ام. دبلیو. براون و آر. کادک^{۹۳} (۱۹۸۹) درباره شاخص‌های اعتبار متقاطع تک نمونه‌ای برای ساختارهای کواریانس طرح شده است.
- معیار اطلاع بیز^{۹۴} (BIC) نیز اولین بار توسط جی. شوارتز^{۹۵} (۱۹۷۸) با نام برآورد بعد یک مدل ارائه شده است. فرمول این شاخص به‌شرح زیر می‌باشد

$$BIC = \chi_m^2 + q(\ln n)$$

- نگارش سازگار معیار اطلاع آکائیک^{۹۶} (CAIC) اولین بار توسط ایچ. بزدوگان^{۹۷} (۱۹۸۷) درباره انتخاب مدل و معیار اطلاع آکائیک مطرح شد. فرمول این شاخص به‌شرح زیر می‌باشد

$$CAIC = \chi_m^2 + q(\ln n + 1)$$

مقدار q در فرمول‌های ارائه شده برابر با تعداد پارامترهای آزاد تعریف شده در مدل تدوین شده است.

^{۹۱} Akaike, H.

^{۹۲} Browne-Cudeck Criterion

^{۹۳} M. W. Brown and R. Codec

^{۹۴} Bayes Information Criterion

^{۹۵} Schwartz, J.

^{۹۶} Consistent Version of Akaike Information Criterion

^{۹۷} Bazdogan, H.

هیچ یک از شاخص‌های ذکر شده نقطه‌ی مشخصی برای قابل قبول تلقی شدن مدل ندارند اما به‌طور کلی می‌توان گفت که هرچه مقدار این شاخص‌ها به صفر نزدیک‌تر شود حاکی از قابل قبول‌تر بودن مدل است. به‌طور کلی هرچه مقدار این شاخص‌ها کوچکتر باشد حاکی از مقتصدتر بودن مدل تفسیر می‌شود. در مجموع این معیارها برای مقایسه مدل‌ها به کار می‌روند و مقدار تکی (منفرد) آن‌ها برای یک مدل به خودی خود تفسیر نمی‌شود. توجه کنید که دو معیار BIC و BCC جریمه بیشتری را برای پیچیدگی مدل در مقایسه با دو شاخص AIC و CAIC قائل می‌شوند.

بیرن (۲۰۰۱) تاکید می‌کند که دو شاخص AIC و CAIC در چارچوب مفهومی مشترکی قرار دارند و هر دوی آن‌ها اندازه‌ای را منعکس می‌کنند که برآورد پارامترها از نمونه اصلی می‌تواند در نمونه‌های دیگر از همان جامعه تکرار شود.

پارامتر غیر مرکزی

پارامتر غیر مرکزی (NCP) همان شاخص کای-دو مدل است که به‌خاطر پیچیدگی مدل جریمه می‌شود. با فرض ثابت بودن سایر شرایط هرچه مقدار کای-دو کوچکتر باشد، مدل قابل قبول‌تر تلقی می‌شود. این شاخص باعث کاهش بیشتر در کای-دو در حالت کمتر پیچیده بودن مدل می‌شود. فرمول محاسبه این شاخص به‌شرح زیر می‌باشد.

$$NCP = Max(\chi_m^2 - df_m, 0)$$

به این معنا که از بین دو مقدار $\chi_m^2 - df_m$ و صفر مقداری که بزرگتر باشد به‌عنوان پارامتر غیر مرکزی در نظر گرفته می‌شود. در مجموع هر چه این پارامتر به صفر نزدیکتر شود مدل قابل قبول‌تر تلقی می‌شود.

شاخص هلتر

این شاخص اساساً با سایر شاخص‌های برازش که تا کنون طرح شد متفاوت بوده و مستقیماً بر موضوع کافی بودن حجم نمونه تمرکز دارد در واقع طرح این شاخص به دنبال تلاشی بوده است که برای یافتن یک شاخص برازش مستقل از حجم نمونه صورت گرفته است. اگر مدل تدوین شده را ثابت در نظر بگیریم و هم‌چنین واریانس‌ها و کواریانس‌ها در ماتریس تغییر داده نشوند با تغییر حجم نمونه شاخص هلتر هیچ تغییری نخواهد کرد هر چند مقدار کای-دو مدل تدوین شده تغییر کند.

هلتر^{۹۸} (۱۹۸۳) مقدار ۲۰۰ را به‌عنوان حداقل قابل قبول برای مقادیر بحرانی N پیشنهاد کرده است. بر اساس همین پیشنهاد مقادیر کوچکتر یا مساوی ۷۵ برای مقدار بحرانی N حاکی از یک برازش ضعیف است. به لحاظ کاربردی می‌توان دو شرط بر اساس شاخص هلتر به‌منظور

قابل قبول بودن مدل و کافی بودن حجم نمونه در یک مطالعه مشخص کرد: اول این که مقدار بحرانی هلتر بیشتر از ۲۰۰ باشد و دوم این که حجم نمونه مورد مطالعه برابر یا بزرگتر از مقدار گزارش شده برای شاخص هلتر باشد. فرمول محاسبه شاخص هلتر یا مقدار بحرانی N به شرح زیر است

$$Hoelter's \quad C.N. = \frac{(n-1) \times \chi_c^2}{\chi_m^2} + 1$$

که در آن مقدار χ_c^2 مقدار بحرانی کای-دو برای آلفای مشخص و درجه آزادی مدل است.

شاخص اعتبار متقاطع

شاخص اعتبار متقاطع مورد انتظار^{۹۹} این احتمال را مورد بررسی قرار می دهد که نتایج حاصل از یک نمونه به نمونه های دیگر با حجم نمونه مشابه تا چه اندازه قابل تکرار است. این شاخص که تنها با یک نمونه قابل محاسبه است بر مبنای تقسیم تصادفی نمونه مورد مطالعه به دو نمونه معیار و نمونه اعتبار قرار دارد. منطق محاسبه آن تفاوت بین ماتریس کواریانس های بازسازی شده حاصل از نمونه مورد مطالعه و ماتریس کواریانس مورد انتظار است که از یک نمونه دیگر با حجم نمونه یکسان از آن جامعه حاصل شده باشد.

از این شاخص می توان برای مقایسه مدل ها با یکدیگر استفاده کرد به نحوی که مقدار کمتر حاکی از قابلیت بیشتر مدل در تکرار نتایج و بنابراین اعتبار بیشتر است. از آن جایی که این شاخص می تواند هر مقداری به خود بگیرد هیچ دامنه مشخصی برای قابل قبول بودن آن تعریف نمی شود. حداقل انتظاری که داریم این است که مقدار ECVI برای مدل تدوین شده در مقایسه با مدل های اشباع شده و استقلال کوچکتر باشد. شاخص اعتبار متقاطع مورد انتظار اصلاح شده^{۱۰۰} مثل BCC است به جز آن که برای عامل مقیاس است، این شاخص برای مقایسه مدل های غیر سلسله مراتبی با داده های یکسان به کار می رود. فرمول محاسبه دو شاخص ذکر شده به شرح زیر است

$$ECVI = \left(\frac{1}{n-1}\right) \times AIC$$

$$MECVI = \left(\frac{1}{n-1}\right) \times BCC$$

اصلاح مدل

اگر برازش یک مدل نظری به قوتی که انتظار داشتیم نبود آن گاه گام بعدی اصلاح مدل و ارزیابی مدل اصلاح شده می باشد. فرآیند نمایان سازی خطاهای تدوین مدل به نحوی که مدل های جایگزین تدوین شده به طور مناسب تری ارزیابی شوند، «جستجوی تدوین» نامیده

^{۹۹}ECVI: Expected cross-validation index

^{۱۰۰}MECVI: Modified expected cross-validation index

می‌شود. هدف از یک جستجوی تدوین تعویض مدل اصلی با مدلی است که در برخی جهات دارای برآزش بهتری بوده و پارامترهایی را برآورد می‌کند که به لحاظ آماری معنادار و به لحاظ نظری دارای معنا و مفهوم باشند. بررسی ماتریس باقیمانده‌ها، ملاحظه معناداری آماری پارامترهای مدل و همچنین استفاده از ضریب لاگرانژ و آماره والد از جمله روش‌های مورد استفاده برای این منظور هستند.

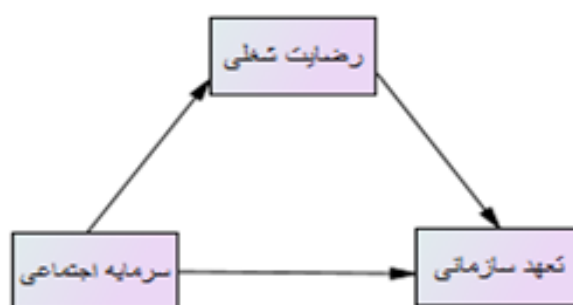
۸.۱ مثال‌های کاربردی

در این جا برای درک بهتر روش SEM به دو مثال کاربردی در حوزه مدیریت ورزشی و روان‌شناسی می‌پردازیم.

۱.۸.۱ مثال مدیریت ورزشی

تاثیر سرمایه اجتماعی و رضایت شغلی بر تعهد سازمانی کارکنان ادارات ورزش و جوانان استان اصفهان و آرایه مدل

با توجه به گستردگی ابعاد فعالیت سازمان‌های ورزشی و نقش مهم آن‌ها در دستیابی به اهداف ورزشی اعم از سطوح قهرمانی، همگانی، تفریحی و سلامت عمومی، لزوم توجه به افزایش تعهد سازمانی و رضایت شغلی کارکنان این سازمان‌ها را بیش از پیش نمایان می‌سازد، زیرا تعهد سازمانی و رضایت شغلی کارکنان یکی از مسائل جدی و غیر قابل انکار در سازمان‌ها به‌شمار می‌رود و عدم توجه به این مسئله منجر به مشکلات عدیده‌ای برای سازمان‌ها خواهد شد اما با این وجود، در سازمان‌ها و مؤسسات توجه کافی بدان مبذول نمی‌شود. از طرفی دیگر با توجه به اینکه در ادارات ورزش و جوانان فعالیت‌های گروهی دارای اهمیت می‌باشد، به‌منظور رشد کارکنان آموزش‌دیده؛ باید محیطی برای ایجاد اعتماد و همکاری بین کارکنان فراهم شود. از این رو با توجه به نکات ذکر شده و با در نظرگیری پژوهش‌های محدود انجام شده در زمینه‌ی بررسی این الگوی ارتباطی در بین کارکنان بالاخص کارکنان ادارات ورزش و جوانان در قالب یک مدل (شکل ۱۳.۱)، محقق در پی بررسی تاثیر سرمایه اجتماعی و رضایت شغلی بر تعهد سازمانی کارکنان ادارات ورزش و جوانان استان اصفهان می‌باشد.



شکل ۱۳.۱: مدل مفهومی پژوهش

شکل ۱۳.۱ مدل مفهومی پژوهش ما را نشان می دهد.

پژوهش حاضر، از نوع پژوهش‌های توصیفی-همبستگی است. به لحاظ هدف، کاربردی و از نظر جمع‌آوری داده‌ها، میدانی است. جامعه آماری این تحقیق را کلیه کارکنان ادارات ورزش و جوانان استان اصفهان تشکیل دادند که بر اساس آمار ارائه شده توسط اداره کل ۲۸۴ نفر اعلام شد. نمونه آماری بر اساس جدول مورگان ۱۶۴ نفر محاسبه شد، که به منظور افزایش صحت و دقت نتایج و احتمال عدم پاسخگویی برخی از نمونه‌ها، ۱۸۵ پرسشنامه توزیع و در نهایت ۱۷۳ پرسشنامه کامل و بدون نقص مورد تجزیه و تحلیل قرار گرفت. نمونه‌های پژوهش به روش تصادفی-خوشه‌ای انتخاب شدند. برای گردآوری اطلاعات از پرسشنامه‌های ویژگی‌های جمعیت شناختی، پرسشنامه ۱۸ سؤالی تعهد سازمانی در قالب سه بعد تعهد عاطفی، مستمر و هنجاری (آلن و مییر، ۱۹۹۰)، پرسشنامه ۲۸ سؤالی سرمایه اجتماعی (ناهایت و قوشال^{۱۰۱}، ۱۹۹۸) در قالب هفت بعد اعتماد، شبکه‌ها، همیاری، روابط، ارزش‌ها، فهم متقابل و تعهد و پرسشنامه سه سؤالی رضایت شغلی میشیگان (کامان، فیچمن، جنکینس و کلش^{۱۰۲}، ۱۹۷۹) استفاده شد. برای سنجش روایی سازه، از تحلیل عاملی تأییدی و برای تعیین روایی محتوایی این پرسشنامه‌ها از نظرات ۱۰ تن از اساتید صاحب‌نظر در این زمینه استفاده شد. هم‌چنین یک مطالعه مقدماتی به منظور تعیین پایایی بر روی ۳۰ نفر از کارکنان ادارات ورزش و جوانان استان اصفهان انجام گرفت که نتایج آلفای کرونباخ برای پرسشنامه سرمایه اجتماعی (۰/۹۰)، پرسشنامه تعهد سازمانی (۰/۸۸) و پرسشنامه رضایت شغلی (۰/۷۵) گزارش شد. برای توصیف داده‌ها از آمار توصیفی و برای تحلیل استنباطی داده‌ها از آزمون ضریب همبستگی پیرسون و هم‌چنین برای تعیین روابط میان متغیرها در قالب مدل علی از مدل سازی معادلات ساختاری بهره گرفته شده است. در انجام آزمون‌های فوق از نرم‌افزارهای آماری Spss23 و Amos23 استفاده شد.

^{۱۰۱}Nahapiet and Ghoshal

^{۱۰۲}Cammann, Fichman, Jenkins and Klesh

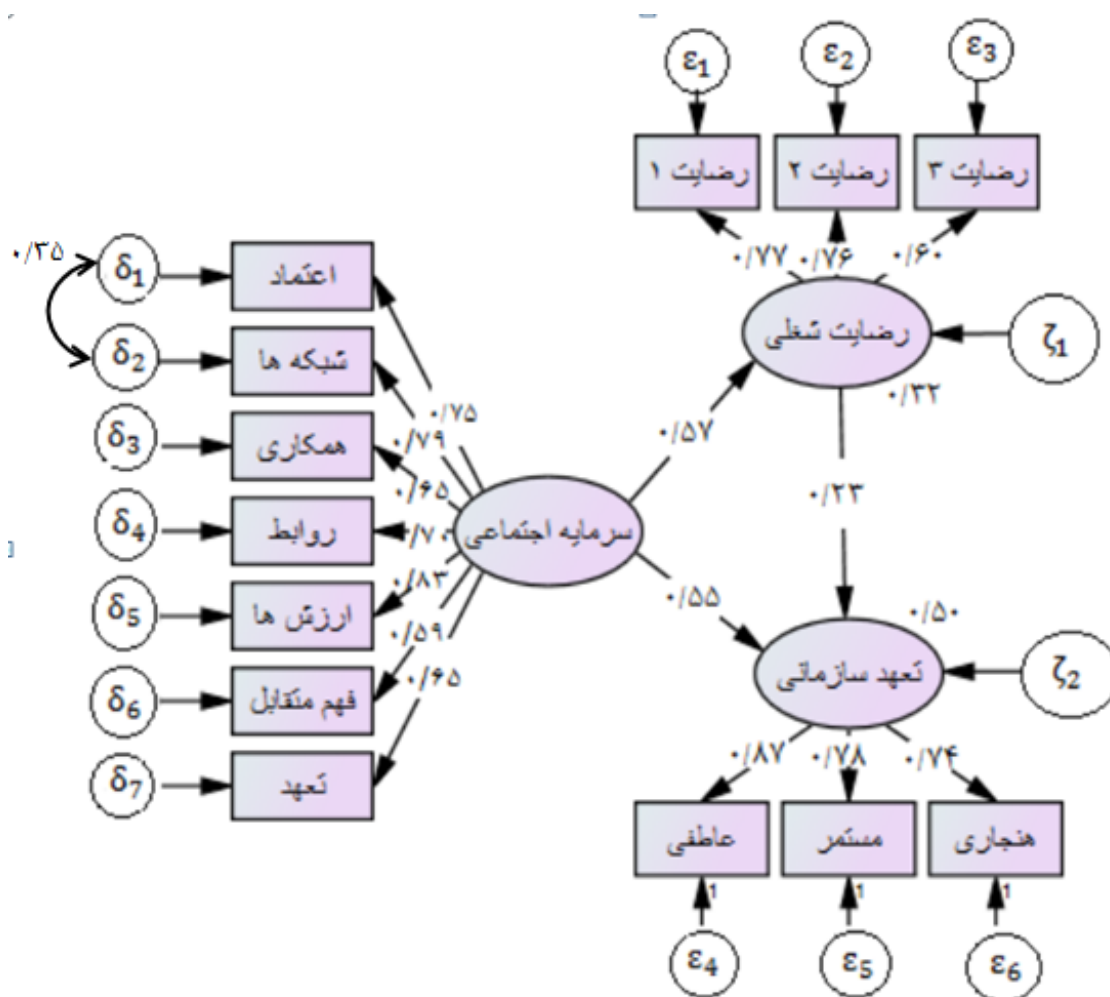
برای تعیین نرمال بودن چندمتغیره از ضریب مردیا^{۱۰۳} استفاده شد و با توجه به نسبت‌های بحرانی به دست آمده ($CR > ۲/۵۸$) فرض عدم نرمال بودن توزیع داده‌ها رد شد. همچنین نتایج مربوط به آماره تلورانس و عامل تورم واریانس نشان داد که بین متغیرهای پیش‌بین همخطی چندگانه وجود ندارد. از آنجایی که مبنای تجزیه و تحلیل مدل‌سازی معادلات ساختاری ماتریس همبستگی است؛ بنابراین ماتریس همبستگی متغیرهای پژوهش در جدول ۱.۱ ارائه شده است.

جدول ۱.۱: ماتریس همبستگی متغیرهای تحقیق

متغیرها	۱	۲	۳
سرمایه اجتماعی	۱		
رضایت شغلی	۰/۴۷**	۱	
تعهد سازمانی	۰/۵۹**	۰/۴۷**	۱

$$p \leq ۰/۰۱**$$

همبستگی تنها درجه رابطه میان سازه‌ها را مشخص می‌کند و به منظور درک بهتر آثار مستقیم و غیرمستقیم و همچنین اثر متغیر میانجی در میان سازه‌ها باید از مدل معادلات ساختاری بهره گرفت.



شکل ۱۴.۱: مدل‌سازی معادلات ساختاری

شکل ۱۴.۱ مدل‌سازی معادلات ساختاری مثال ما را نشان می‌دهد.

در مدل معادلات ساختاری از یک طرف میزان انطباق داده‌های پژوهش و مدل مفهومی پژوهش بررسی خواهد شد که آیا از برازش مناسب برخوردار است و از طرف دیگر معناداری روابط در این مدل برازش یافته، آزمون می‌شود. با توجه به عدم توافق عمومی و کلی در بین متخصصان مدلیابی معادلات ساختاری بر روی بهترین شاخص‌های برازندگی برآورد، در تحقیق حاضر از بین شاخص‌های برازندگی مطلق، شاخص χ^2/df (CMIN/DF)، شاخص نیکویی برازش (GFI) و شاخص ریشه میانگین مجذور برآورد تقریب (RMSEA) و از بین شاخص‌های برازندگی تطبیقی، شاخص برازندگی تطبیقی (CFI)، شاخص توکر-لوپس (TLI) و شاخص برازش هنجار شده (NFI) مورد استفاده قرار گرفت که توسط میرز^{۱۰۴} و همکاران (۲۰۱۲) پیشنهاد شده است.

^{۱۰۴}Meyers

جدول ۲.۱: شاخص‌های برازش تأثیر متغیر سرمایه اجتماعی و تعهد سازمانی بر رضایت شغلی

CMIN	RMSEA	CMIN/DF	NFI	CFI	TLI
۱۰۶/۹۶	۰/۰۶۶	۱/۷۵	۰/۹۰	۰/۹۵	۰/۹۴

جدول ۲.۱ مقادیر شاخص‌های برازش تأثیر متغیر سرمایه اجتماعی و تعهد سازمانی بر رضایت شغلی را نشان می‌دهد.

زمانی که $TLI, CFI, NFI > ۰/۹$ ، $CMIN/DF < ۳$ و $RMSEA < ۰/۰۸$ برازش مدل خوب است. با توجه به ستون مقادیر بحرانی در جدول زیر، مشخص می‌شود که هر سه مسیر دارای تأثیر معنادار $|t| \geq ۱/۹۶$ می‌باشند که در جدول ۲.۱ نمایش داده شده‌اند. تحلیل مدل ساختاری نشان داد که سرمایه اجتماعی تأثیر مستقیم و معناداری بر تعهد سازمانی و رضایت شغلی دارد. همچنین تأثیر معنادار رضایت شغلی و تعهد سازمانی مورد تأیید قرار گرفت.

جدول ۳.۱: ضرایب استاندارد، انحراف استاندارد و مقدار بحرانی متغیرهای پژوهش

متغیرها	ضریب استاندارد	انحراف استاندارد	مقدار بحرانی
سرمایه اجتماعی —> تعهد سازمانی	۰/۵۵	۰/۱۰	۵/۰۵
رضایت شغلی —> تعهد سازمانی	۰/۲۳	۰/۰۶	۲/۲۸
سرمایه اجتماعی —> رضایت شغلی	۰/۵۷	۰/۱۴	۵/۶۶

در جدول زیر ابتدا اثرات مستقیم سپس اثرات غیرمستقیم و اثرات کل مورد بحث و بررسی قرار می‌گیرد. همان‌گونه که در جدول ۴.۱ نیز مشاهده می‌شود، سرمایه اجتماعی متغیری است که هم به صورت مستقیم و هم به صورت غیرمستقیم و به واسطه متغیر رضایت شغلی بر تعهد سازمانی تأثیر معناداری دارد. متغیر سرمایه اجتماعی بر رضایت شغلی و رضایت شغلی بر تعهد سازمانی نیز تأثیر مستقیم و معناداری دارد. نتایج مندرج در ستون اثرات کل نشان داد که سرمایه اجتماعی بیشترین تأثیر را بر رضایت شغلی دارد. در مدل پژوهش ۱۴.۱ مشاهده می‌شود، ۳۲ درصد از واریانس متغیر رضایت شغلی توسط سرمایه اجتماعی و ۵۰ درصد از واریانس تعهد سازمانی توسط رضایت شغلی و سرمایه اجتماعی قابل تبیین می‌باشد؛

کارکنان می‌توانند در کار خود نظم بیشتری داشته باشند، مدت زمان بیشتری را در سازمان صرف کنند و در راستای رسیدن به اهداف سازمان از هیچ تلاش و کوششی مضایقه ننمایند، این‌ها از مصادیق تعهد سازمانی می‌باشند که کارکنان نسبت به سازمان خود وفادار و از خود گذشتگی نشان می‌دهند (کاظم زاده بیطالی و قاسم زاده علیشاهی، ۱۳۹۲). یافته‌های پژوهش حاضر با نتایج ملک مکان و همکاران (۲۰۱۵)، صیادی (۲۰۱۴)، امام و همکاران (۲۰۱۴)، مایک و همکاران (۲۰۱۲)، علمی و همکاران (۱۳۹۰) و فخرایی و اسدی (۱۳۸۷) هم‌راستا می‌باشد.

در ادامه نتایج حاصل از تحلیل مدل پژوهش نشان داد که رضایت شغلی اثر مستقیم و معناداری بر تعهد سازمانی دارد. بر اساس نظریه برابری می‌توان ادعان نمود که علی‌رغم علاقه درونی یا ذاتی فرد نسبت به کار خود اگر پاداش یا تنبیه در محیط کار بر پایه عدل و انصاف یا شایسته‌سالاری ارائه شود، کارکنان از شغل خود احساس رضایت می‌کنند؛ حال هر چه این احساس رضایت قوی‌تر باشد، میل وی برای پیوسته ماندن به سازمان بیشتر می‌شود و به مرور زمان وابستگی عاطفی به سازمان پیدا کرده و خود را متعهد به ادامه کار در این سازمان می‌داند (حیدری تفریشی و دریابگیان، ۱۳۹۰). از این‌رو چنانچه سازمان‌ها فرصت‌هایی را به منظور افزایش رضایت کارکنان خود فراهم نمایند می‌توانند سطح تعهد کارکنان خود را نیز ارتقاء دهند تا بدین وسیله سبب بهبود کارایی و اثربخشی شوند. یافته‌های پژوهش حاضر با نتایج پژوهش‌های نات گانگای و آگراوال (۲۰۱۵)، سریواستوا (۲۰۱۳)، سوما و لسا (۲۰۱۳) و جهانگیر و همکاران (۱۳۸۶) همخوانی دارد.

در نهایت یافته‌های پژوهش حاضر نشان داد که سرمایه اجتماعی تأثیری مستقیم و معنادار بر رضایت شغلی دارد. نظر به این‌که وزارت ورزش و جوانان، متولی اصلی ورزش کشور محسوب می‌شود و این وظیفه خطیر در جامعه ما بر عهده کارکنان ادارات ورزش و جوانان استان‌هاست، بایستی در تأمین رفاه اجتماعی آنان کوشید. بی‌چون و چرا میزان رضایت کارکنان آثار خود را در نظام ورزش که در پرورش توانایی‌های نسل جوان و توسعه و تعمیم ورزش تأثیر بسزایی دارد، نشان خواهد داد. یکی از عوامل بسیار مهم دوام و بقاء هر سازمانی، شبکه روابط و کنش متقابل بین افراد تشکیل‌دهنده آن سازمان هستند. حال اگر این روابط مبتنی بر مؤلفه‌هایی باشند که در این پژوهش به‌عنوان سرمایه اجتماعی از آن نام برده شده، می‌توان محیطی خوشایند و ارضاء‌کننده جهت اعضای سازمان به‌وجود آورد و در نهایت به افزایش بهره‌وری و توسعه سازمان یاری رسانید. به‌طور کلی با داشتن شبکه روابط اجتماعی قوی از طریق روابط با همکاران، اطمینان داشتن از وعده‌های همکاران، فراهم کردن فرصت‌هایی برای عضویت در گروه‌های اجتماعی و رفع و رجوع مشکلات از طریق کار گروهی مؤثر، می‌توان در کارکنان رضایت بیشتری ایجاد کرد. این نتایج با یافته‌های پژوهش چمنی‌فرد و نیک‌پور (۲۰۱۵)، سواری و همکاران (۲۰۱۳)، عبدالملکی و همکاران (۱۳۹۳)، حقیقتیان و همکاران (۱۳۹۰) و امن و همکاران (۲۰۰۹) هم‌راستا می‌باشد.

۲.۸.۱ مثال روان‌شناسی

مثال ثبات از خود بیگانگی

مطالعه‌ای را در نظر بگیرید که در ارتباط با ثبات نگرش یا طرز برخورد در طول زمان مانند خود بیگانگی و تاثیر و ارتباط این نگرش بر متغیرهای دیگری شبیه تحصیل و اشتغال است. بدین منظور داده‌هایی در مقیاس نگرش از ۹۳۲ نفر در دو ناحیه روستایی در ایلینوس و در سه نقطه زمانی ۱۹۶۶، ۱۹۶۷ و ۱۹۷۱ جمع‌آوری شده‌اند. در اینجا تنها داده‌های مربوط به ۱۹۶۷ و ۱۹۷۱ را بررسی می‌کنیم. مقادیر مقیاس کم خونی^{۱۰۵} (آنمی) و مقیاس ضعف^{۱۰۶} به‌عنوان نمایندگان (نشان‌هایی از) متغیر پنهان مفروض یعنی از خود بیگانگی در نظر گرفته می‌شوند. هم‌چنین فرض شده است که سال‌های تحصیل پرسش شونده (مخاطب) و شاخص جامعه اقتصادی دانکن^{۱۰۷} نمایندگان یک حالت اجتماعی اقتصادی^{۱۰۸} مخاطب باشند.

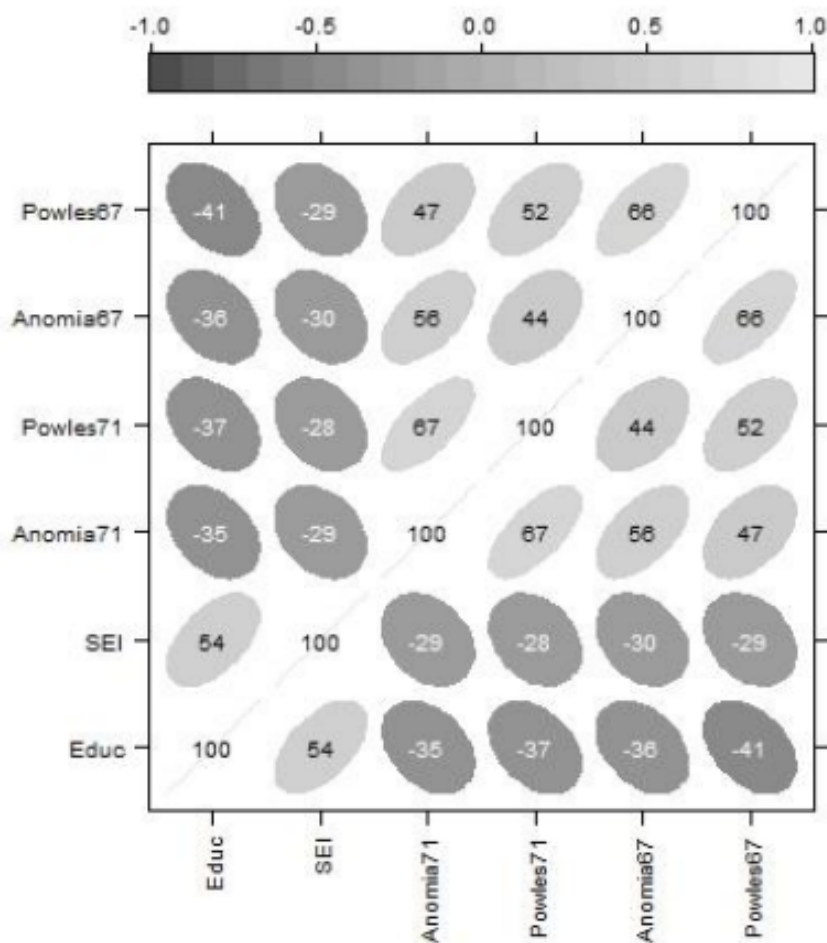
ماتریس همبستگی برای شش متغیر مشاهده شده در شکل زیر با استفاده از این کد نشان داده شده است.

^{۱۰۵}Anemia

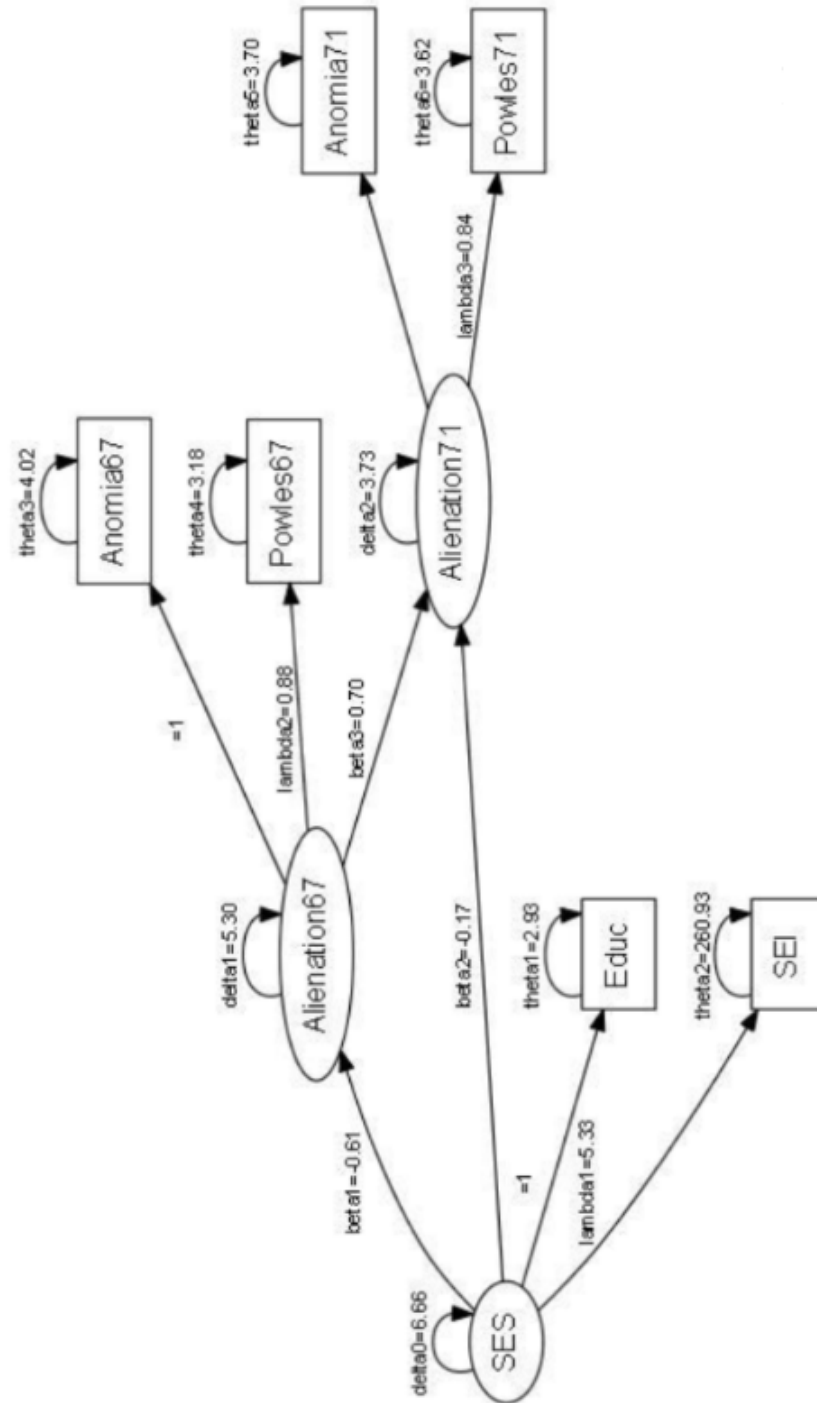
^{۱۰۶}Powerlessness

^{۱۰۷}Duncan's socioeconomic index

^{۱۰۸}Socioeconomic status



نمودار مسیر برای مدلی که برازش شده در شکل زیر نشان داده شده است. فرض شده است که متغیر پنهان حالت جامعه اقتصادی در هر دو نقطه زمانی بر از خود بیگانگی تاثیر دارد و همچنین از خود بیگانگی در سال ۱۹۶۷ بر از خود بیگانگی در سال ۱۹۷۱ تاثیر می‌گذارد.



مقیاس سه متغیر پنهان SES، از خود بیگانگی ۵۷ و از خود بیگانگی ۷۱ دلخواه هستند و باید به طریقی برای قابل شناسایی بودن مدل یک در نظر گرفته شوند. در نتیجه معادلات مدل برازش شده به صورت زیر می باشند.

$$\begin{aligned}
 Education &= SES + u_1 \\
 SEI &= \lambda_1 SES + u_2 \\
 Anomia_{67} &= Alienation_{67} + u_3 \\
 Powerlessness_{67} &= \lambda_2 Alienation_{67} + u_4 \\
 Anomia_{71} &= Alienation_{71} + u_5 \\
 Powerlessness_{71} &= \lambda_3 Alienation_{71} + u_6 \\
 Alienation_{67} &= \beta_1 SES + u_7 \\
 Alienation_{71} &= \beta_2 SES + \beta_3 Alienation_{67} + u_8
 \end{aligned}$$

علاوه بر شش ضریب رگرسیون در این معادلات، واریانس‌های هشت عبارت خطا u_1, u_2, \dots, u_8 و واریانس عبارت خطا برای متغیرهای پنهان SES نیز باید برآورد شود. کد لازم R برای برازش مدل در زیر آورده شده است.

```

alienation_model<- specify.model( file = "alienation_model.txt")
alienation_sem<- sem(alienation_model, alienation, 932)

```

که مدل زیر را در نظر می‌گیرد.

```

SES ->Educ, NA, 1
SES -> SEI, lambda1, NA
Alienation67 -> Anomia67, NA, 1
Alienation67 -> Powles67, lambda2, NA
Alienation71 -> Anomia71, NA, 1
Alienation71 -> Powles71, lambda3, NA
SES -> Alienation67, beta1, NA
SES -> Alienation71, beta2, NA
Alienation67 -> Alienation71, beta3, NA
Educ<->Educ, theta1, NA
SEI <-> SEI, theta2, NA
SES <-> SES, delta0, NA
Anomia67 <-> Anomia67, theta3, NA
Powles67 <-> Powles67, theta4, NA
Anomia71 <-> Anomia71, theta5, NA

```

Powles71 <-> Powles71, theta6, NA
 Alienation67 <-> Alienation67, delta1, NA
 Alienation71 <-> Alienation71, delta2, NA

برآورد پارامترها عبارتند از

summary(alienation_sem)

Model Chisquare= 71.532 Df = 6 Pr(>Chisq) = 1.9825e-13
 Chisquare (null model) = 2131.5 Df = 15
 Goodness-of-fit index = 0.97514
 Adjusted goodness-of-fit index = 0.913
 RMSEA index = 0.10831 90% CI: (0.086636, 0.1315)
 Bentler-Bonnett NFI = 0.96644
 Tucker-Lewis NNFI = 0.9226
 Bentler CFI = 0.96904
 SRMR = 0.021256
 AIC = 101.53
 AICc= 72.056
 BIC = 174.09
 CAIC = 24.508

Normalized Residuals

Min.1st Qu.Median Mean 3rd Qu. Max.

-1.26000 -0.20900 -0.00001 -0.01510 0.24400 1.33000

R-square for Endogenous Variables

Educ SEI Alienation67 Anomia67 Powles67 Alienation71

0.6942 0.4205 0.3210 0.6602 0.6594 0.5764

Anomia71 Powles71

0.7045 0.6372

Parameter Estimates

Estimate Std Error z value Pr(>|z|)

```
lambda1 5.33054 0.429754 12.4037 2.4959e-35 SEI <--- SES
lambda2 0.88882 0.041530 21.4020 1.2788e-101 Powles67 <--- Alienation67
lambda3 0.84892 0.039965 21.2417 3.9310e-100 Powles71 <--- Alienation71
beta1 -0.61362 0.056441 -10.8718 1.5701e-27 Alienation67 <--- SES
beta2 -0.1744 0.053897 -3.2370 1.2078e-03 Alienation71 <--- SES
beta3 0.70463 0.053532 13.1629 1.4346e-39 Alienation71 <--- Alienation67
theta1 2.93613 0.499371 5.8797 4.1112e-09 Educ<-->Educ
theta2 260.93180 18.238709 14.3065 1.9934e-46 SEI <--> SEI
delta0 6.66387 0.640557 10.4032 2.3963e-25 SES <--> SES
theta3 4.02303 0.343389 11.7157 1.0596e-31 Anomia67 <--> Anomia67
theta4 3.18936 0.271559 11.7446 7.5264e-32 Powles67<--> Powles67
theta5 3.70313 0.373384 9.9178 3.4850e-23 Anomia71 <--> Anomia71
theta6 3.62335 0.292079 12.4054 2.4430e-35 Powles71 <--> Powles71
delta1 5.30687 0.472629 11.2284 2.9577e-29 Alienation67 <--> Alienation67
delta2 3.73998 0.387452 9.6528 4.7851e-22 Alienation71 <--> Alienation71
```

Iterations = 86

مقدار آماره کای-دو با ۶ درجه آزادی برابر با ۷۱/۵۳۲ است که نشان می‌دهد مدل به‌خوبی برازش نشده است پیشنهاد شده است که می‌توان با در نظر گرفتن این که خطاهای اندازه‌گیری برای آنمی در ۱۹۶۷ و ۱۹۷۱ همبسته‌اند مدل را بهبود بخشید. برای برازش چنین مدلی باید خط زیر را به کد بالا اضافه کرد.

```
Anomia67 <-> Anomia71, psi, NA
```

با کد زیر

```
mod2 <- c(mod, "Anomia67 <-> Anomia71,psi,NA")
writeLines(mod2, con = "alienation_model2.txt")
```

Model Chisquare= 6.3587Df = 5 Pr(>Chisq) = 0.27286

Chisquare (null model) = 2131.5Df = 15

Goodness-of-fit index = 0.99772

Adjusted goodness-of-fit index = 0.99043

RMSEA index = 0.017085 90% CI: (NA, 0.051008)

Bentler-Bonnett NFI = 0.99702

Tucker-Lewis NNFI = 0.99807

Bentler CFI = 0.99936

SRMR = 0.0108

AIC = 38.359

AICc= 6.9532

BIC = 115.76

CAIC = -32.828

Normalized Residuals

Min.1st Qu.Median Mean 3rd Qu. Max.

-0.9600 -0.1340 -0.0429 -0.0914 0.0000 0.5450

R-square for Endogenous Variables

Educ SEI Alienation67 Anomia67 Powles67 Alienation71

0.7166 0.4073 0.3063 0.5721 0.7638 0.5010

Anomia71 Powles71

0.6168 0.7313

Parameter Estimates

Estimate Std Error z value Pr(>|z|)

lambda1 5.16348 0.421283 12.2566 1.5495e-34 SEI <--- SES

lambda2 1.02693 0.053185 19.3088 4.5284e-83 Powles67 <--- Alienation67

lambda3 0.97091 0.049411 19.6496 5.8287e-86 Powles71 <--- Alienation71

beta1 -0.54948 0.053396 -10.2906 7.7721e-25 Alienation67 <--- SES

beta2 -0.21173 0.049249 -4.2993 1.7137e-05 Alienation71 <--- SES

beta3 0.61727 0.049694 12.4216 1.9953e-35 Alienation71 <--- Alienation67

theta1 2.72054 0.515876 5.2736 1.3375e-07 Educ<-->Educ

theta2 266.86570 18.189641 14.6713 9.8439e-49 SEI <--> SEI

مثال‌های کاربردی ۵۷

```
delta0 6.87946 0.657502 10.4630 1.2771e-25 SES <--> SES
theta3 5.07293 0.371338 13.6612 1.7310e-42 Anomia67 <--> Anomia67
theta4 2.21167 0.317868 6.9578 3.4557e-12 Powles67 <--> Powles67
theta5 4.81198 0.395167 12.1771 4.1172e-34 Anomia71 <--> Anomia71
theta6 2.68355 0.329805 8.1368 4.0593e-16 Powles71 <--> Powles71
delta1 4.70500 0.433171 10.8618 1.7535e-27 Alienation67 <--> Alienation67
delta2 3.86587 0.343460 11.2556 2.1724e-29 Alienation71 <--> Alienation71
psi 1.88823 0.240076 7.8651 3.6868e-15 Anomia71 <--> Anomia67
```

Iterations = 89

حال آماره‌ی کای-دو با ۵ درجه آزادی برابر با ۶/۳۵۹ است. بدیهی است داخل کردن خطاهای اندازه‌گیری همبسته در مدل برای دو اندازه‌آنی برازش مدل را بسیار بهبود داده است.

فصل ۲

مدل معادلات ساختاری ریج

۱.۲ مقدمه

در این فصل، مدل معادلات ساختاری ریج را معرفی کرده به برخی خواص مجانبی مهم آن از جمله سازگاری و نرمال بودن توزیع برآورد پارامتر اشاره می‌کنیم. به عنوان پیش درآمدی بر مدل ریج، ابتدا به‌طور مختصر مشکل همخطی در مدل‌های رگرسیونی را عنوان کرده سپس برآوردگر ریج بردار ضرایب مدل رگرسیون خطی چندگانه را ارائه می‌کنیم. دلیل این یادآوری از رگرسیون ریج این است که می‌خواهیم از ایده مطرح شده در ارائه مدل معادلات ساختاری ریج استفاده کنیم. به‌طور کلی موضوعات همخطی و رگرسیون ریج برگرفته از کتاب مونت گمری (۲۰۱۲) و مدل معادلات ساختاری ریج از مقاله یوان و همکاران (۲۰۱۱) می‌باشد.

۲.۲ همخطی در رگرسیون چندگانه

غالباً تفسیر و استفاده از الگوی رگرسیون چندگانه به برآورد تک‌تک ضرایب رگرسیون بستگی دارد. وقتی هیچ رابطه خطی میان متغیرهای پیش‌بین وجود نداشته باشد، این متغیرها متعامد^۱ نامیده می‌شوند. هنگامی که متغیرهای پیش‌بین متعامد باشند، استنباط در خصوص

^۱Orthogonal

پارامترهای مدل به طور معمول (روش کمترین توان‌های دوم) انجام می‌شود. اما در اکثر کاربردهای رگرسیونی متغیرهای پیش‌بین متعامد نیستند. اگر در یک الگوی رگرسیون متغیرهای پیش‌بین به طور کامل یا تقریباً کامل، وابسته خطی باشند، در این حالت ممکن است استنباط برای این الگوی رگرسیون گمراه کننده باشد. در این صورت می‌گوییم بین متغیرها در الگوی رگرسیون، همخطی^۲ وجود دارد. اگر یکی از متغیرهای پیش‌بین، تابع دقیق خطی از یک یا چند متغیر پیش‌بین دیگر باشد، گوییم رگرسیون دارای همخطی کامل^۳ است. همخطی ناقص^۴ وقتی اتفاق می‌افتد که یکی از متغیرهای پیش‌بین به طور تقریبی یک تابع خطی از یک یا چند متغیر پیش‌بین دیگر باشد. به عبارت دیگر همخطی کامل زمانی رخ می‌دهد که دست‌کم به ازای یک i ($i = 1, 2, \dots, p$)، $R_i^2 = 1$ باشد و همخطی ناقص زمانی پیش می‌آید که لااقل به ازای یک i ، $R_i^2 \cong 1$ است، که در آن $R_i^2 = 1$ ضریب تعیین^۵ رگرسیون خطی متغیر x_i بر سایر متغیرهای پیش‌بین است. برای اطلاع از چگونگی ایجاد همخطی کامل و ناقص به روزه (۱۳۹۰) مراجعه کنید.

اما همخطی اثراتی بر برآورد پارامترها دارد که با توجه به این اثرات نتیجه می‌گیریم روش‌های معمول استنباط مناسب نیستند. همخطی ناقص باعث می‌شود که دترمینان ماتریس $X^T X$ بسیار کوچک شده و در نتیجه ماتریس $(X^T X)^{-1}$ دارای اعضایی باشد که به لحاظ قدرمطلق بسیار بزرگ است و به بی‌نهایت میل کنند. به عبارت دیگر، واریانس برآوردگر کمترین توان‌های دوم β یعنی $\hat{\beta} = (X^T X)^{-1} X^T Y$ بسیار بزرگ می‌شود و در این صورت، برآوردهای به دست آمده بزرگتر از مقدار واقعی برآورد شده و ممکن است علامت‌های اشتباه داشته باشند. به عبارت دیگر دقت برآورد ضرایب رگرسیون بسیار کم می‌شود. علاوه بر این، $\hat{\beta}$ ناپایدار است. یعنی با تغییر کوچکی در مقادیر متغیرهای پیش‌بین، مقادیر موجود در $\hat{\beta}$ تغییرات زیادی پیدا می‌کنند و در نتیجه محاسبه $\hat{\beta}$ با دو کامپیوتر متفاوت (با دوروش متفاوت) ممکن است جواب‌های کاملاً متفاوتی به دست دهد. از طرفی چون $Var(\hat{\beta}) = \sigma^2 (X^T X)^{-1}$ ، همخطی باعث به دست آمدن فواصل اطمینان عریض برای پارامتر β می‌شود. همچنین در صورت وجود همخطی کامل، ماتریس $X^T X$ وارون‌پذیر نیست و در نتیجه برای برآورد ضرایب، جواب یکتا وجود نخواهد داشت.

بنابراین، باید در هر مسئله رگرسیونی ابتدا وجود همخطی بررسی شده و راهکاری مناسب برای رویارویی با مشکلات ذکر شده در نظر گرفته شود. به طور معمول، سه روش عمده برای تشخیص همخطی در مدل رگرسیون خطی چندگانه وجود دارد.

محک ماتریس همبستگی

فرض کنید X ماتریس $n \times p$ باشد. محک بسیار ساده همخطی چندگانه، بررسی اعضای

^۲ Collinearity

^۳ Exact collinearity

^۴ Partial collinearity

^۵ Coefficient of determination

غیرقطری r_{ij} به ازای $i, j = 1, 2, \dots, p$ در ماتریس $X^T X$ است. اگر متغیرهای پیش بین X_i و X_j تقریباً وابسته خطی باشند، در این صورت $|r_{ij}|$ نزدیک یک خواهد بود.

عامل تورم واریانس

یکی از شیوه‌های تشخیص وجود همخطی که کاربرد زیادی دارد، عامل تورم واریانس^۶ (VIF) است. این عامل نشان می‌دهد که واریانس ضرایب برآورد شده تا چه حد نسبت به حالتی که متغیرهای پیش بین همبستگی خطی ندارند، متورم شده است. عامل تورم واریانس از رابطه زیر به دست می‌آید

$$VIF_j = \frac{1}{1 - R_j^2} \quad j = 1, 2, \dots, p-1$$

که در آن، R_j^2 ضریب تعیین مدلی است که در آن متغیر مستقل X_j روی سایر متغیرهای مستقل برازش داده شده است. در حالتی که $R_j^2 = 0$ یا به عبارت دیگر X_j با سایر متغیرهای مستقل در مدل رگرسیونی رابطه خطی نداشته باشد، $VIF_j = 1$ و هنگامی که $R_j^2 \neq 0$ ، VIF_j بزرگتر از ۱ خواهد بود که نشان می‌دهد واریانس X_j به دلیل وجود همبستگی بین متغیرهای پیش بین چه قدر متورم شده و مقدار آن افزایش پیدا کرده است. هنگامی که یک متغیر پیش بین، وابستگی خطی کامل با سایر متغیرهای پیش بین مدل داشته باشد، ضریب تعیین مرتبط با آن برابر با یک خواهد بود. در نتیجه عامل تورم واریانس به سمت بی‌نهایت میل خواهد کرد و از آن می‌توان نتیجه گرفت که واریانس X_j نیز بی‌نهایت خواهد شد. مقدار عامل تورم واریانس برای متغیرهای پیش بین، اغلب به عنوان شاخصی برای اندازه‌گیری شدت همخطی در مدل استفاده می‌شود. اگر بیشترین مقدار عامل تورم واریانس در مدل بزرگتر از ۱۰ باشد، نشان‌دهنده این است که وجود همخطی در مدل، اثر نامطلوبی بر برآوردهای روش کمترین توان‌های دوم گذاشته است.

عدد شرطی

ساده‌ترین راه برای پی بردن به وجود همخطی به دست آوردن عدد شرطی^۷ ماتریس $X^T X$ است. برای محاسبه عدد شرطی ابتدا باید مقادیر ویژه ماتریس $X^T X$ را به دست آوریم. مقدار عدد شرطی از نسبت بزرگترین مقدار ویژه به کوچکترین مقدار ویژه حاصل می‌شود.

$$C.N. = \frac{\lambda_{Max}}{\lambda_{min}} \quad (1.2)$$

در حالت کلی اگر عدد شرطی کمتر از ۱۰۰ باشد مشکل جدی در مورد همخطی چندگانه وجود ندارد. چنانچه این مقدار بزرگتر از ۱۰۰ باشد همخطی شدید و اگر از ۱۰۰۰ تجاوز کند همخطی خیلی جدی است.

لازم به ذکر است که در بعضی مقالات جذر رابطه (۱.۲) به عنوان تعریف عدد شرطی در نظر گرفته می‌شود است و در این صورت اگر جذر عدد شرطی $X^T X$ بزرگتر از ۱۰ باشد، آن گاه در

^۶Variance inflation factor

^۷Condition number

ماتریس طرح همخطی وجود دارد. اگر جذر عدد شرطی بین 3° تا 100° ، همخطی شدید و بالاتر از 100° همخطی کاملاً جدی و مشکل ساز است. برای آگاهی بیشتر در خصوص محک‌های ذکر شده به مونت گمری (۲۰۱۲) مراجعه کنید.

۳.۲ برآوردگر ریج

در بخش قبل گفتیم که وقوع همخطی چندگانه بین متغیرهای پیش‌بین در تحلیل رگرسیون خطی ممکن است باعث ناپایداری شدید در برآورد کمترین توان‌های دوم پارامترها شود. در نتیجه، برآورد کمترین توان‌های دوم به‌دست آمده، قابل اعتماد نخواهد بود.

روش‌های متعددی برای غلبه بر مشکل همخطی در مدل‌های خطی ارائه شده است. اما مسئله‌ی مهمی باید در استفاده از این برآوردگرها در نظر گرفت این است که برخلاف برآوردگر کمترین توان‌های دوم، این روش‌ها منجر به ایجاد برآوردی اریب می‌شوند. در عین حال معمولاً این رهیافت‌ها به پارامتر مجهول وابسته هستند که می‌توان با انتخاب درست این پارامتر، با علم به این که برآوردگر اریب است، مقدار میانگین توان دوم خطا را کاهش داد. برخی از این روش‌ها عبارتند از: برآورد انقباضی^۸، مؤلفه‌های اصلی^۹، ریج^{۱۰}، کمترین توان‌های دوم جزئی^{۱۱} و نوع لیو^{۱۲}. برای اطلاع از این تکنیک‌ها می‌توانید به آکدنیز و کاجیرانلار^{۱۳} (۱۹۹۵)، سوئیندل^{۱۴} (۱۹۷۶) و لیو^{۱۵} (۲۰۰۳) مراجعه کنید. در این پایان‌نامه، ما از روش ریج استفاده می‌کنیم.

یکی از مؤثرترین روش‌ها برای حل مشکل همخطی، برآوردگر ریج است که توسط هورل و کنارد در سال ۱۹۷۰ معرفی شد. مدل رگرسیون خطی (۲.۱) را در نظر بگیرید، برآوردگر ریج به پارامتر ریج k بستگی دارد که با افزودن این پارامتر به ماتریس $X^T X$ باعث بهبود وضعیت آن می‌شود

$$\begin{aligned}\hat{\beta}(k) &= (X^T X + kI_p)^{-1} X^T Y \\ &= (I_p + k(X^T X)^{-1})^{-1} (X^T X)^{-1} X^T Y \\ &= T(k)\hat{\beta}\end{aligned}$$

که در آن $T(k) = (k(X^T X)^{-1} + I_p)^{-1}$. لازم به ذکر است در فصل ۳، برآوردگر ریج از دیدگاه مدل‌های جریمه شده توضیح داده شده است. در واقع برآوردگر ریج یک تبدیل خطی از برآوردگر

^۸ Shrinkage estimation

^۹ Principal components

^{۱۰} Ridge

^{۱۱} Partial least squares

^{۱۲} Liu type

^{۱۳} Akdeniz and Kaciranlar

^{۱۴} Swindel

^{۱۵} Liu

کمترین توان‌های دوم است. امید ریاضی $\hat{\beta}(k)$ عبارتست از

$$E(\hat{\beta}(k)) = E(\mathbf{T}(k)\hat{\beta}) = \mathbf{T}(k)E(\hat{\beta}) = \mathbf{T}(k)\beta$$

بنابراین، $\hat{\beta}(k)$ برآوردگر اریب β است. هم‌چنین ماتریس واریانس-کواریانس آن به صورت زیر می‌باشد

$$\begin{aligned} \text{Cov}(\hat{\beta}(k)) &= \text{Cov}((\mathbf{X}^T\mathbf{X} + k\mathbf{I}_p)^{-1}\mathbf{X}^T\mathbf{Y}) \\ &= (\mathbf{X}^T\mathbf{X} + k\mathbf{I}_p)^{-1}\mathbf{X}^T\text{Cov}(\mathbf{Y})\mathbf{X}(\mathbf{X}^T\mathbf{X} + k\mathbf{I}_p)^{-1} \\ &= \sigma^2(\mathbf{X}^T\mathbf{X} + k\mathbf{I}_p)^{-1}\mathbf{X}^T\mathbf{X}(\mathbf{X}^T\mathbf{X} + k\mathbf{I}_p)^{-1} \end{aligned}$$

در بخش بعد از ایده طرح شده در بالا یعنی جایگذاری $(\mathbf{X}^T\mathbf{X})$ با $(\mathbf{X}^T\mathbf{X} + k\mathbf{I}_p)$ برای برطرف کردن مشکل همخطی استفاده کرده و مدل معادلات ساختاری ریج را معرفی می‌کنیم. برای آگاهی بیشتر درخصوص مشکل همخطی و برآورد ریج می‌توان به پایان‌نامه‌های فارسی حسن‌زاده بشتیان (۱۳۸۸)، نجاریان (۱۳۹۰)، برزوئی بیدگلی (۱۳۹۳)، آرست (۱۳۹۵) و رساله دکتری روزبه (۱۳۹۰) مراجعه کنید.

۴.۲ مدل معادلات ساختاری ریج

در مطالعات مدل‌سازی معادلات ساختاری روش‌هایی که در حال حاضر در نرم‌افزارها در دسترس هستند روش‌هایی دو مرحله‌ای می‌باشند که همبستگی‌های چندحالتی و چندرشته‌ای را در قالب ماتریس همبستگی مدل‌بندی می‌کنیم. این ماتریس همبستگی با استفاده از روش‌های ماکزیمم درست‌نمایی و مجموع توان‌های دوم خطی تعمیم یافته مجانبی و کمترین توان‌های دوم قطری وزن‌دار که در نرم‌افزارهای EQS، LISREL و Mplus موجود می‌باشند، برآورد می‌شود. برای آگاهی بیشتر در خصوص این روش‌ها بابکاس و همکاران (۱۹۸۷)، دولان (۱۹۹۴)، دیستفانو (۲۰۰۲) و لیو (۲۰۰۹) را ببینید.

اخیراً لیو (۲۰۰۹) روش ماکزیمم درست‌نمایی را در نرم‌افزار EQS و روش کمترین توان‌های دوم قطری وزن‌دار را در نرم‌افزار Mplus مطالعه کرد و دریافت که روش دوم نرخ همگرایی بهتری نسبت به روش اول دارد. او هم‌چنین دریافت که اریبی نسبی برآورد پارامترهای این دو روش تقریباً مشابه هستند. به‌طور خلاصه این دو روش، روش‌هایی هستند که به‌طور گسترده‌ای مورد مطالعه قرار می‌گیرند.

هنگامی که ماتریس واریانس-کواریانس نمونه S تقریباً یکنواخت است، در این صورت ماتریس $S_k = S + kI$ با یک مقدار عددی مثبت k ، می‌تواند جایگزین مناسبی باشد. یوان و چان (۲۰۰۸) مدل S_k را به‌جای مدل S پیشنهاد کردند. آن‌ها نشان دادند که نتایج این روش در برآورد پارامترها سازگار است. نتایج تجربی نشان می‌دهد که نه تنها همگرایی بهتری دارد، بلکه در حجم نمونه کوچک نتایج برآورد پارامترها دقیق‌تر از برآوردگر ماکزیمم درست‌نمایی،

حتی زمانی که توزیع داده‌ها نرمال می‌باشد، است. در این فصل روش یوان و همکاران (۲۰۱۱) را مورد بررسی قرار می‌دهیم.

هدف این است که روش یوان و چان (۲۰۰۸) را به SEM با داده‌های رسته‌ای و پیوسته با مدل‌سازی ماتریس همبستگی چندحالتی، چندرشته‌ای و گشتاور- حاصلضربی گسترش دهیم. نرم‌افزار LISREL گزینه‌ای برای مدل‌سازی $S = kdiag(s_{11}, \dots, s_{pp})$ فراهم می‌کند که گزینه‌ی ریج نامیده می‌شود. به راحتی در حالتی که بخواهیم ماتریس همبستگی R را مدل کنیم گزینه‌ی ریج در LISREL را به $R + kI$ برازش می‌دهیم که همان گسترش روش یوان و چان (۲۰۰۸) برای ماتریس‌های همبستگی می‌باشد. روش ریج با داده‌های رسته‌ای در حال حاضر در نرم‌افزار LISREL اجرا شده است. اگرچه به علت فقدان خاصیت‌های آن واضح نیست که چگونه این روش در عمل به درستی اجرا می‌شود.

در واقع مک کوییتی (۱۹۹۷) مطالعات تجربی را بر روی گزینه‌ی ریج در نسخه ۸ نرم‌افزار LISREL انجام داد و به این نتیجه رسید که به نظر می‌رسد شواهد کافی وجود دارد مبنی بر این که LISREL با گزینه‌ی ریج در LISREL تخمین زده نمی‌شود وجود دارد، مگر این که تنها هدف، برآورد غیر استاندارد بارهای عاملی باشد.

یوان و همکاران (۲۰۱۱) خواص آماری SEM ریج با داده‌های رسته‌ای را به دست آوردند و آن را به صورت یک روش آماری دقیق (بی‌عیب) ارائه کردند. هم‌چنین نشان دادند که SEM ریج با داده‌های رسته‌ای دارای برآورد پارامتر و خطای استاندارد سازگار است و چهار آماره را برای ارزیابی مدل کلی پیشنهاد کردند.

فرض کنید R یک ماتریس همبستگی $p \times p$ ، شامل همبستگی‌های چندحالتی، چندرشته‌ای و گشتاور- حاصلضربی پیرسون برای متغیرهای رسته‌ای و پیوسته باشد. فرض کنید r برداری از همه‌ی شکل‌های همبستگی با عناصر پایین قطری R باشد و ρ همانند جامعه r باشد. سپس آن از جورسکوگ (۱۹۹۴)، لی و همکاران (۱۹۹۵) یا موتن و ستورا (۱۹۹۵) که

$$\sqrt{n}(r - \rho) \xrightarrow{D} N(0, \Upsilon). \quad (2.2)$$

پیروی می‌کند که در آن Υ ماتریس کواریانس مجانبی از \sqrt{nr} است که می‌تواند بدون تناقض برآورد شود. بر اساس محک تابع درست‌نمایی می‌توان معیار اختلاف زیر را نسبت به θ کمینه کرد.

$$F_{ML}(\theta) = tr[R\Sigma^{-1}(\theta)] - \log |R\Sigma^{-1}(\theta)| - p \quad (3.2)$$

در این صورت برآوردگر ماکزیمم درست‌نمایی θ عبارتست از

$$\hat{\theta} = \underset{\theta}{\operatorname{argmin}} F_{ML}(\theta) \quad (4.2)$$

نرم‌افزار SEM گزینه‌ای دارد که این کمینه‌سازی را بر حسب θ انجام داده و برآوردگر $\hat{\theta}$ را برای مدل‌بندی R نتیجه می‌دهد.

حال فرض کنید، $R_k = R + kI$ که در آن $k > 0$ است. به جای کمینه کردن رابطه‌ی (۳.۲) SEM، عبارت

$$F_{ML}(\theta_k) = tr[R_k \Sigma_k^{-1}(\theta_k)] - \log |R_k \Sigma_k^{-1}(\theta_k)| - p \quad (۵.۲)$$

را برای برآورد پارامتر $\hat{\theta}_k$ کمینه می‌کند، که در آن $\Sigma_k(\theta_k) = \Sigma(\theta_k) + kI$ می‌باشد.

۵.۲ خواص برآوردگر ریج در مدل معادلات ساختاری

۱.۵.۲ سازگاری

فرض کنید $z = (x^{*T}, y^{*T})^T$ متغیرهای یک مدل SEM باشد، مدل اندازه‌گیری آن را به صورت زیر در نظر بگیرید.

$$\begin{cases} x^* = \mu_x + \Lambda_x \xi + \delta \\ y^* = \mu_y + \Lambda_y \xi + \varepsilon \end{cases} \quad (۶.۲)$$

که در آن $\mu_x = E(x^*)$ ، $\mu_y = E(y^*)$ ، Λ_x و Λ_y ماتریس بارهای عاملی می‌باشند؛ ξ و η بردارهایی از ساختار پنهان با $E(\xi) = 0$ و $E(\eta) = 0$ ؛ و δ و ε بردارهایی از خطاهای اندازه‌گیری با $E(\delta) = 0$ ، $E(\varepsilon) = 0$ ، $E(\delta\delta^T) = \Theta_\delta$ و $E(\varepsilon\varepsilon^T) = \Theta_\varepsilon$ هستند. همچنین مدل ساختاری که روابط بین ξ و η را شرح می‌دهد را به صورت زیر در نظر بگیرید

$$\eta = B\eta + \Gamma\xi + \zeta$$

که در آن ζ برداری از خطاهای پیش‌بینی با $E(\zeta) = 0$ و $\Psi = E(\zeta\zeta^T)$ می‌باشد. فرض کنید که $\Phi = E(\xi\xi^T)$ باشد، در این صورت ساختار کواریانس z برابر است با

$$\Sigma(\theta) = \begin{pmatrix} \Lambda_x \Phi \Lambda_y + \Theta_\delta & \Lambda_x \Phi \Gamma^T (I - B')^{-1} \Lambda_y^T \\ \Lambda_y (I - B)^{-1} \Gamma \Phi \Lambda_x^T & \Lambda_y (I - B)^{-1} (\Gamma \Phi \Gamma^T + \Psi) (I - B)^{-1} \Lambda_y^T + \Theta_\varepsilon \end{pmatrix}. \quad (۷.۲)$$

اکنون فرض کنید $\mu_y = 0$ ، $\mu_x = 0$ و $diag(\Theta_\delta) = I - diag(\Lambda_x \Phi \Lambda_y)$ و $diag(\Theta_\varepsilon) = I_\varepsilon - diag[\Lambda_y (I - B)^{-1} (\Gamma \Phi \Gamma^T + \Psi) (I - B^T)^{-1} \Lambda_y^T]$ که در آن $diag(A)$ ماتریس قطری از عناصر قطری A ، و I_δ و I_ε ماتریس همانی، به ترتیب از بعد Θ_δ و Θ_ε می‌باشند. بنابراین عناصر قطری Θ_δ و Θ_ε بخشی از پارامترهای آزاد نیستند اما بخشی از مدل $\Sigma(\theta)$ از میان توابع پارامترهای آزاد در Λ_y ، Λ_x ، B ، Γ ، Φ ، Ψ ، عناصر غیر قطری (Θ_δ) و عناصر غیر قطری (Θ_ε) است.

زمانی $\Sigma(\theta)$ برای Σ مدل درستی است که ماتریس‌های $\Lambda_x^{(0)}$ ، $\Lambda_y^{(0)}$ ، $B^{(0)}$ ، $\Gamma^{(0)}$ ، $\Phi^{(0)}$ ، $\Psi^{(0)}$ ، $offdiag(\Theta_\delta^{(0)})$ و $offdiag(\Theta_\varepsilon^{(0)})$ وجود داشته باشند به طوری که $\Sigma = \Sigma(\theta_0)$ باشد، که در آن θ_0 برداری شامل مقادیر جامعه از همه‌ی پارامترهای آزاد در θ می‌باشد.

فرض کنید θ_k مربوط به برداری برای $\Lambda_x^{(k)} = \Lambda_x^{(o)}$ ، $\Lambda_y^{(k)} = \Lambda_y^{(o)}$ ، $B^{(k)} = B^{(o)}$ ، $\Gamma^{(k)} = \Gamma^{(o)}$ ، $\Phi^{(k)} = \Phi^{(o)}$ ، $\Psi^{(k)} = \Psi^{(o)}$ ، $offdiag(\Theta_\delta^{(k)}) = offdiag(\Theta_\delta^{(o)})$ و $offdiag(\Theta_\varepsilon^{(k)}) = offdiag(\Theta_\varepsilon^{(o)})$ باشد. علاوه بر این فرض کنید

$$diag(\Theta_\delta^{(k)}) = diag(\Theta_\delta^{(o)}) + kI_\delta \quad (۸.۲)$$

و

$$diag(\Theta_\varepsilon^{(k)}) = diag(\Theta_\varepsilon^{(o)}) + kI_\varepsilon \quad (۹.۲)$$

بنابراین، $\theta^{(k)} = \theta^{(o)}$ ، $\Theta_\delta^{(k)}$ و $\Theta_\varepsilon^{(k)}$ توابعی از $\theta^{(o)}$ و k هستند. فرض کنید توابع موجود در رابطه (۸.۲) و (۹.۲) بخشی از مدل $\Sigma_k = \Sigma_k(\theta_k)$ باشند. توجه کنید که $\Sigma_k(\theta)$ به طور منحصر به فردی با $\Sigma(\theta)$ و k تعیین شده است. هرگاه $\Sigma(\theta)$ مدل درستی برای مدل سازی Σ باشد، $\Sigma_k(\theta)$ مدل درستی برای مدل سازی Σ_k می باشد. هم چنین نتایج نشان می دهد که به ازای $\Sigma(\theta)$ خاص، مدل سازی R_k ، پارامترهای Λ_x ، Λ_y ، B ، Γ ، Φ ، Ψ و نیز عناصر قطری غیرقطری Θ_δ و Θ_ε را مشابه مدل سازی R برآورد می کند. اما برآورد عناصر قطری ماتریس های Θ_δ و Θ_ε به ازای مقادیر مختلف k ثابت دلخواه متفاوت است. معمولاً، عناصر قطری $\hat{\Theta}_\delta$ و $\hat{\Theta}_\varepsilon$ با واریانس خطاهای اندازه گیری به طور منحصر به فردی برآورد می شوند. زمانی که R_k را مدل سازی می کنیم، آن ها را می توان از طریق روابط زیر به دست آورد.

$$diag(\hat{\Theta}_\delta) = diag(\hat{\Theta}_\delta^{(k)}) - kI_\delta. \quad (۱۰.۲)$$

$$diag(\hat{\Theta}_\varepsilon) = diag(\hat{\Theta}_\varepsilon^{(k)}) - kI_\varepsilon. \quad (۱۱.۲)$$

بحث فوق بیان گر آن است که، به دلیل وجود خطای نمونه گیری ممکن است مدل سازی R_k ، برآورد $\hat{\theta}_k$ را متفاوت از برآورد $\hat{\theta}$ در مدل سازی R نتیجه دهد.

۲.۵.۲ نرمال مجانبی

یوان و همکاران^{۱۶} (۲۰۱۱) توزیع مجانبی $\hat{\theta}_k$ را به دست آوردند، و در نتیجه سازگاری SEs را به دست می آوریم. برای این منظور ابتدا برخی از نمادها را معرفی می کنیم. برای ماتریس A ، فرض کنید $vech(A)$ برداری باشد که از زیر هم نوشته شدن ستون های ماتریس پایین مثلثی A به دست می آید.

اکنون $s_k = vech(R_k)$ ، $\sigma_k(\theta) = vech[\Sigma_k(\theta)]$ و $s = vech(R)$ ، $\sigma(\theta) = vech[\Sigma(\theta)]$ را تعریف می کنیم. توجه کنید که تفاوت بین r و s_k سبب می شود، s_k شامل p عنصر $k+1$ روی

قطر R_k باشد. فرض کنید D_p تکرار ماتریس تعریف شده توسط مگ نوس و نئودکر^{۱۷} (۱۹۹۹) باشد و

$$W_k(\theta) = \frac{1}{\gamma} D_p^T [\Sigma_k^{-1}(\theta) \otimes \Sigma_k^{-1}(\theta)] D_p.$$

که در آن \otimes ضرب کرونگر^{۱۸} می باشد. به عنوان مثال، اگر θ شامل q پارامتر مجهول باشد، $\sigma_k(\theta) = \frac{\partial \sigma_k(\theta)}{\partial \theta^T}$ یک ماتریس $p^* \times q$ است که در آن $p^* = \frac{p(p+1)}{\gamma}$ می باشد. تحت شرایط نظم استاندارد، فرض کنید θ_0 یک نقطه داخلی از Θ باشد و $\hat{\theta}_k$ در رابطه‌ی زیر صدق کند.

$$g_k(\hat{\theta}_k) = 0 \quad (12.2)$$

که در آن

$$g_k(\theta) = \sigma_k^T(\theta) W_k(\theta) [s_k - \sigma_k(\theta)]$$

است. در رابطه‌ی (۱۲.۲) از آن جا که $\sigma(\theta)$ و $\sigma_k(\theta)$ تنها تفاوتشان در مقدار ثابت k است، $\sigma_k(\theta) = \sigma(\theta)$ و $s_k - \sigma_k(\theta) = s - \sigma(\theta)$ برقرار می باشند. بنابراین تاثیر a روی $\hat{\theta}_k$ در رابطه‌ی (۱۲.۲) تنها با $W_k(\theta)$ است که

$$W_k(\theta) = \frac{1}{\gamma} D_p^T [\Sigma(\theta) + kI]^{-1} \otimes [\Sigma(\theta) + kI]^{-1} D_p = \frac{1}{\gamma(k+1)^2} U_k \quad (13.2)$$

که در آن

$$U_k = D_p^T \left[\frac{1}{k+1} \Sigma(\theta) + \frac{k}{k+1} I \right]^{-1} \otimes \left[\frac{1}{k+1} \Sigma(\theta) + \frac{k}{k+1} I \right]^{-1} D_p.$$

توجه داشته باشید که $\frac{1}{\gamma(k+1)^2}$ در رابطه (۱۳.۲) ضریب ثابت است و هیچ تاثیری بر روی $\hat{\theta}_k$ ندارد. U_k این تفاوت را ایجاد می کند. زمانی که $k = 0$ و $\hat{\theta}_k = \hat{\theta}$ برآورد پارامتر به روش ML صورت می گیرد (همان روش ML است). زمانی که $k = \infty$ ،

$$U_k = D_p^T D_p$$

ماتریس وزنی به دست آمده از برآوردگر LS ، ماتریس همبستگی R را مدل سازی می کند. بنابراین SEM ریج می تواند به عنوان ترکیبی از روش های ML و LS در نظر گرفته شود. انتظار داریم که مزایای هر دو روش را داشته باشد، یعنی SEM ریج نرخ همگرایی بهتر و دقیق تری نسبت به برآورد پارامتر ML و نیز برآوردهای کاراتری در مقایسه با برآوردگر LS داشته باشد.

بر اساس رابطه (۱۲.۲) و بسط تیلور $g_k(\hat{\theta})$ در θ_0 می توان نوشت

$$\sqrt{n}(\hat{\theta}_k - \theta_0) = -[g_k(\bar{\theta})]^{-1} g_k(\theta_0) = (\sigma_k^T W_k \sigma_k)^{-1} \sigma_k^T W_k \sqrt{n}(s_k - \sigma_k) + o_p(1). \quad (14.2)$$

^{۱۷}Meg Nuss and Nojodker

^{۱۸}kroncker Multiply

که در آن $g_k(\bar{\theta})$ یک ماتریس $q \times q$ و هر سطر آن در یک بردار $\bar{\theta}$ که بین θ_0 و $\hat{\theta}_k$ است ارزیابی می‌شود و $o_p(1)$ نشان‌دهنده‌ی مقداری است که با افزایش n در احتمال همگرا به صفر است. توجه کنید سطرهای σ_k مربوط به عناصر قطری Σ_k ، صفر و همچنین بردار $(s_k - \sigma_k)$ و $(r - \rho)$ بردارهایی هستند که از جمع p تا عدد صفر تشکیل شده‌اند. با توجه به رابطه (۲.۲)

$$\sqrt{n}(s_k - \sigma_k) \xrightarrow{D} N(o, \Upsilon^*) \quad (15.2)$$

که در آن Υ^* ماتریس $p^* \times p^*$ است که به صورت زیر تعریف می‌شود. ممکن است که رابطه (۱۵.۲) را با تعریف کلی از یک متغیر تصادفی درک نکنیم، که فقط یک ثابت زمانی است که واریانس آن صفر باشد. طبق رابطه (۱۴.۲) و (۱۵.۲)

$$\sqrt{n}(\hat{\theta}_k - \theta_0) \xrightarrow{D} N(o, \Omega) \quad (16.2)$$

که در آن

$$\Omega = (\sigma_k^T W_k \sigma_k)^{-1} \sigma_k^T W_k \Upsilon^* W_k \sigma_k (\sigma_k^T W_k \sigma_k). \quad (17.2)$$

یک برآورد ثابت از $\hat{\Omega}$ از Ω را زمانی که پارامترهای مجهول (نامعلوم) در رابطه (۱۷.۲) با $\hat{\theta}_k$ و Υ^* توسط $\hat{\Omega}^*$ جایگذاری کنیم، به دست آورد. توجه کنید Ω در رابطه (۱۷.۲) ماتریس کواریانس مجانبی می‌باشد.

۳.۵.۲ آماره‌هایی برای ارزیابی کلی مدل

این زیر بخش چهار آماره برای ارزیابی کلی مدل ارائه می‌کند. هنگامی که رابطه (۵.۲) را برای برآورد پارامتر کمینه می‌کنیم، به طور خودکار اختلاف بین داده‌ها و مدل را اندازه می‌گیریم، به عنوان مثال $F_{MLk}(\hat{\theta}_k)$. اگرچه آماره معروف $T_{MLk} = nF_{MLk}(\hat{\theta}_k)$ مجانباً از توزیع کای-دو حتی زمانی که $k = 0$ است پیروی نمی‌کند. فرض کنید

$$F_{RLS_k}(\hat{\theta}_k) = [s_k - \sigma_k(\hat{\theta}_k)]^T W_k(\hat{\theta}_k) [s_k - \sigma_k(\hat{\theta}_k)] = \frac{1}{p} tr [R_k \Sigma_k^{-1}(\hat{\theta}_k) - I]^2$$

و

$$T_{RLS_k} = nF_{RLS_k}(\hat{\theta}_k) \quad (18.2)$$

به اصطلاح آماره LS بازبینی می‌شود، در خروجی EQS به طور پیش فرض ماتریس کواریانس را مدل سازی می‌کند. تحت فرض ساختار مدل صحیح، رابطه زیر برقرار است.

$$T_{MLk} = T_{RLS_k} + o_p(1). \quad (19.2)$$

بر طبق رابطه (۱۴.۲) داریم

$$\begin{aligned} \sqrt{n}[\mathbf{s}_k - \boldsymbol{\sigma}_k(\hat{\boldsymbol{\theta}}_k)] &= \sqrt{n}(\mathbf{s}_k - \boldsymbol{\sigma}_k) - [\boldsymbol{\sigma}_k(\hat{\boldsymbol{\theta}}_k) - \boldsymbol{\sigma}_k(\boldsymbol{\theta}_0)] & (۲۰.۲) \\ &= \sqrt{n}(\mathbf{s}_k - \boldsymbol{\sigma}_k) - \dot{\boldsymbol{\sigma}}_k(\hat{\boldsymbol{\theta}}_k - \boldsymbol{\theta}_0) + o_p(1) \\ &= \mathbf{P}_k \sqrt{n}(\mathbf{s}_k - \boldsymbol{\sigma}_k) + o_p(1), \end{aligned}$$

که در آن

$$\mathbf{P} = \mathbf{I} - \boldsymbol{\sigma}_k(\boldsymbol{\sigma}_k^T \mathbf{W}_k \boldsymbol{\sigma}_k)^{-1} \boldsymbol{\sigma}_k^T \mathbf{W}_k.$$

با ترکیب رابطه (۱۹.۲) و (۲۰.۲) رابطه زیر به دست می آید.

$$T_{ML_k} = n(\mathbf{s}_k - \boldsymbol{\sigma}_k)^T \mathbf{P}_k^T \mathbf{W}_k \mathbf{P}_k (\mathbf{s}_k - \boldsymbol{\sigma}_k) + o_p(1). \quad (۲۱.۲)$$

توجه کنید که Υ^* در رابطه (۱۵.۲) دارای درجه‌ای از $p_* = \frac{p(p-1)}{q}$ می باشد، یک ماتریس $p_* \times p_*$ مانند A وجود دارد که $A^T A = \Upsilon^*$. فرض کنید $\mathbf{u} \sim N_{p_*}(\mathbf{0}, \mathbf{I})$ ، بر طبق رابطه (۱۵.۲) داریم

$$\sqrt{n}(\mathbf{s}_k - \boldsymbol{\sigma}_k) = \mathbf{A}\mathbf{u} + o_p(1). \quad (۲۲.۲)$$

حاصل ترکیب رابطه (۱۶.۲) و (۱۷.۲) برابر است با

$$T_{ML_k} = \mathbf{u}^T (\mathbf{P}_k \mathbf{A})^T \mathbf{W}_k (\mathbf{P}_k \mathbf{A}) \mathbf{u} + o_p(1). \quad (۲۳.۲)$$

توجه کنید که $(\mathbf{P}_k \mathbf{A})^T \mathbf{W}_k (\mathbf{P}_k \mathbf{A})$ مثبت منتهای است و رتبه آن $p_* - q$ است. فرض کنید $0 < k_1 \leq k_2 \leq \dots \leq k_{p_*-q}$ مقادیر ویژه غیر صفر از $(\mathbf{P}_k \mathbf{A})^T \mathbf{W}_k (\mathbf{P}_k \mathbf{A})$ یا هم‌ارز (معادل) $\mathbf{P}_k^T \mathbf{W}_k \mathbf{P}_k \Upsilon^*$ بر طبق رابطه (۲۳.۲) داریم

$$T_{ML_k} = \sum_{j=1}^{p_*-q} k_j u_j^2 + o_p(1). \quad (۲۴.۲)$$

که در آن u_j^2 ها پیشگو هستند و هر کدام از χ_1^2 پیروی می کنند. مگر این که همه‌ی k_j ها ۱ باشند در این صورت توزیع T_{ML_k} بر طبق $\chi_{p_*-q}^2$ نخواهد بود. هر چند ممکن است رفتار T_{ML_k} تقریباً توسط توزیع کای-دو با همان میانگین توصیف شده باشد. فرض کنید $\hat{\Upsilon}^*$ برآوردگری سازگار Υ^* ، که می تواند از جمع p سطر و p ستون صفر از $\hat{\Upsilon}$ بدست آید، و

$$\hat{m} = \text{tr}(\hat{\Upsilon}^* \hat{\mathbf{P}}_k^T \hat{\mathbf{W}}_k \hat{\mathbf{P}}_k) / (p_* - q).$$

سپس هنگامی که $n \rightarrow \infty$

$$T_{RML_k} = T_{ML_k} / \hat{m}$$

فرض کنید

$$m_1 = \sum_{j=1}^{p_*-q} k_j^2 / \sum_{j=1}^{p_*-q} k_j, \quad m_2 = \left(\sum_{j=1}^{p_*-q} k_j \right)^2 / \sum_{j=1}^{p_*-q} k_j^2$$

سپس T_{ML_k}/m_1 به طور تقریبی با $\chi_{m_2}^2$ در دو لحظه اول برابر است. برآوردهای متناظر m_1 و m_2 به صورت زیر به دست می آیند.

$$\hat{m}_1 = tr[(\hat{Y}^* \hat{P}_k^T \hat{W}_k \hat{P}_k)^2] / tr(\hat{Y}^* \hat{P}_k^T \hat{W}_k \hat{P}_k), \quad (25.2)$$

$$\hat{m}_2 = [tr(\hat{Y}^* \hat{P}_k^T \hat{W}_k \hat{P}_k)]^2 / tr[(\hat{Y}^* \hat{P}_k^T \hat{W}_k \hat{P}_k)^2]$$

بنابراین از تقریب زیر استفاده می کنیم

$$T_{AML_k} = T_{ML_k} / \hat{m}_2 \sim \chi_{\hat{m}_2}^2 \quad (26.2)$$

که ممکن است منجر به توصیف بهتر T_{ML_k} نسبت به رابطه $T_{RML_k} \sim \chi_{p^*-q}^2$ شود. همانند T_{ML_k} ، زمانی که R_k را مدل سازی می کنیم، برطبق توزیع کای-دو حتی زمانی که $k=0$ است، آماره T_{RLS_k} مجانبی نمی باشد. آن از رابطه (۱۹.۲) که توزیع T_{RLS_k} را می توان با استفاده از تقریب به صورت

$$T_{RRLS_k} = T_{RLS_k} / \hat{m} \sim \chi_{p^*-q}^2 \quad (27.2)$$

یا

$$T_{ARLS_k} = T_{RLS_k} / \hat{m}_1 \sim \chi_{\hat{m}_2}^2 \quad (28.2)$$

به دست آورد پیروی می کند.

۶.۲ یک مثال واقعی

ایسنک و ایسنک^{۱۹} (۱۹۷۵) پرسش نامه شخصیتی را طراحی کردند که بعداً در سال ۱۹۸۳ گانگ^{۲۰} آن را به زبان چینی ترجمه کرد. این پرسش نامه در بین ۱۱۷ ورزشکار فارغ التحصیل یک دانشگاه توزیع شد. چهار زیرمقیاس در این پرسش نامه وجود دارد (برون گرایی / درون گرایی، روان رنجوری / پایداری، روان / جامعه پذیری، دروغ) و هر زیرمقیاس شامل ۲۰ تا ۲۴ مورد می باشد که در دو گروه طبقه بندی شده است. به دو بخش داده از زیرمقیاس برون گرایی / درون گرایی دسترسی داریم که ۲۱ موردند. بر اساس کتابچه راهنمای پرسش نامه، این ۲۱ مورد، یک رفتار پنهان (ویژگی نهفته) از مقیاس برون گرایی / درون گرایی پاسخ دهندگان را نشان می دهد. بنابراین ممکن است بخواهیم داده ها را به وسیله یک مدل تک عاملی به دو بخش برآزش دهیم.

جدول ۱ شامل برآورد پارامترها و خطای استاندارد (۲۱ مورد) مقیاس برون گرایی و درون گرایی است زمانی که $R_k = R + kI$ را با مقادیر $k = 0/5, 0/6, 0/7$ مدل می کنیم.

^{۱۹}Eysenck and Eysenck

^{۲۰}Gong

نتایج نشان می‌دهد که هم برآورد پارامتر و هم خطای استاندارد با افزایش k از $۵/۰$ به $۷/۰$ مقدار کمی تغییر می‌کنند.

جدول ۲ آماره‌های T_{ARLS_k} ، T_{RML_k} ، T_{AML_k} ، T_{RRLS_k} همراه با p-value و نیز \hat{m}_2 برای آماره‌های تعدیل‌یافته را نشان می‌دهد.

آماره‌ها نشان می‌دهند که مدل داده‌ها را خوب برازش می‌کند هنگامی که مقدار k زیاد شود مقادیر آماره‌های T_{RML_k} ، T_{AML_k} کاهش می‌یابند. آماره‌های T_{RRLS_k} ، T_{ARLS_k} و هم‌چنین \hat{m}_2 به سختی تحت تاثیر مقدار k قرار می‌گیرند. مقایسه‌ی آماره‌ها نشان می‌دهد که آماره‌های به‌دست آمده از T_{ML_k} ناشی از T_{RLS_k} ممکن است قابل اعتماد نباشد. نتایج این جدول را می‌توان به‌صورت زیر بیان کرد

۱- مقادیر آماره‌ها حاکی از آن است که مدل داده‌ها را خوب برازش می‌کند.

۲- با افزایش مقدار k ، مقادیر آماره‌های T_{RML_k} و T_{AML_k} کاهش می‌یابد.

۳- تاثیر تغییر مقدار k بر آماره‌های T_{RRLS_k} و T_{ARLS_k} و نیز \hat{m}_2 چندان محسوس نیست.

۷.۲ بحث و نتیجه‌گیری

به‌کارگیری روش SEM برای داده‌های رسته‌ای با بیشتر نرم‌افزارها امکان‌پذیر است. با این حال مشکلات همگرایی در برآورد پارامترها و عدم وجود آماره‌های معتبر برای ارزیابی مدل کلی وجود دارد، به‌ویژه هنگامی که اندازه‌ی نمونه کوچک باشد و فراوانی‌های مشاهده شده توزیع اریب داشته باشند. در این‌جا یوان و همکاران (۲۰۱۱) رهیافت ریج جفتی^{۲۱} را با روش ML مطالعه کردند و نشان دادند که برآورد پارامترها سازگار هستند و توزیع نرمال مجانبی دارند و هم‌چنین چهار آماره برای ارزیابی مدل کلی ارائه کردند. نتایج تجربی نشان داد که روش ریج در نرخ همگرایی، سرعت همگرایی، دقت و بهره‌وری از برآورد پارامتر بهتر از روش ML عمل می‌کند و آماره‌ی مقیاس تصویر T_{RRLS_k} بهترین عملکرد را در اندازه نمونه کوچک‌تر دارد. برای SEM با ماتریس کواریانس، یوان و چان^{۲۲} (۲۰۰۸) پیشنهاد کردند که $k = \frac{p}{n}$ انتخاب کنیم. از آن‌جایی که $\frac{p}{n} \rightarrow 0$ و $n \rightarrow \infty$ نتیجه گرفتند که برآوردگرهای مجانبی با برآوردگرهای ML معادل می‌باشند. برخلاف یک ماتریس کواریانس که همیشه نامنفی متناهی است، ماتریس همبستگی چندحالتی، چندرشته‌ای ممکن است مقادیر ویژه منفی که قدر مطلق آن‌ها از $\frac{p}{n}$ بیشتر است اختیار کند. انتخاب $k = \frac{p}{n}$ ممکن است منجر به یک مقدار مثبت متناهی R_k نشود.

یوان و همکاران (۲۰۱۱) فقط روش ML ریج را مطالعه کردند زیرا روش ML محبوب‌ترین روش است و به‌طور گسترده‌ای در SEM مورد استفاده قرار می‌گیرد.

^{۲۱} Pair ridge

^{۲۲} Yuan and chan

با وجود این که کارآمدتر از برآوردهای ML هستند. برآوردهای منفی از واریانس خطا می‌توانند با برآوردهای ریج رخ دهند، این دلیلی است که $\Sigma_k(\theta)$ مبهم می‌باشد. اگر $\Sigma_k(\theta)$ مبهم شدن آن ادامه می‌دهیم که R_k را مدل می‌کنیم. برای درستی مدل‌های مشخص، برآوردهای منفی از واریانس خطا صرفاً به دلیل خطای نمونه‌گیری است، که در هنگام ارزیابی باید کارایی تجربی و اریبی شمارش شوند. زمانی که $\Sigma(\theta)$ و $z \sim N(\mu, \Sigma)$ هر دو مبهم باشند، ممکن است این دو در هم آمیخته نشوند.

جدول ۱.۲: برآورد پارامترها و خطای استاندارد آنها

k	$\hat{\theta}_k$			SE		
	۰/۵	۰/۶	۰/۷	۰/۵	۰/۶	۰/۷
λ_1	۰/۶۶۴	۰/۶۶۳	۰/۶۶۳	۰/۱۱۹	۰/۱۱۹	۰/۱۱۹
λ_2	۰/۷۴۱	۰/۷۴۱	۰/۷۴۲	۰/۰۶۹	۰/۰۷۰	۰/۰۷۰
λ_3	۰/۸۲۷	۰/۸۲۶	۰/۸۲۶	۰/۰۶۳	۰/۰۶۳	۰/۰۶۴
λ_4	۰/۳۰۸	۰/۳۱۰	۰/۳۱۱	۰/۱۲۱	۰/۱۲۱	۰/۱۲۱
λ_5	۰/۷۱۱	۰/۷۱۳	۰/۷۱۴	۰/۰۷۹	۰/۰۷۹	۰/۰۷۹
λ_6	۰/۵۵۱	۰/۵۵۲	۰/۵۵۴	۰/۰۸۹	۰/۰۸۸	۰/۰۸۸
λ_7	-۰/۶۵۳	-۰/۶۵۴	-۰/۶۵۵	۰/۰۸۶	۰/۰۸۵	۰/۰۸۵
λ_8	۰/۶۹۴	۰/۶۹۵	۰/۶۹۵	۰/۰۷۲	۰/۰۷۲	۰/۰۷۲
λ_9	-۰/۶۴۱	-۰/۶۴۳	-۰/۶۴۴	۰/۰۷۸	۰/۰۷۸	۰/۰۷۸
λ_{10}	۰/۸۳۵	۰/۸۳۷	۰/۸۳۸	۰/۰۸۰	۰/۰۸۰	۰/۰۸۰
λ_{11}	۰/۵۴۹	۰/۵۴۹	۰/۵۵۰	۰/۰۹۸	۰/۰۹۸	۰/۰۹۸
λ_{12}	۰/۵۹۵	۰/۵۹۶	۰/۵۹۷	۰/۰۹۵	۰/۰۹۵	۰/۰۹۴
λ_{13}	-۰/۶۵۸	-۰/۶۵۷	-۰/۶۵۷	۰/۰۹۹	۰/۰۹۹	۰/۰۹۹
λ_{14}	۰/۸۱۰	۰/۸۰۹	۰/۸۰۷	۰/۰۶۳	۰/۰۶۴	۰/۰۶۴
λ_{15}	۰/۶۵۱	۰/۶۴۹	۰/۶۴۷	۰/۱۶۰	۰/۱۶۰	۰/۱۶۰
λ_{16}	۰/۳۲۴	۰/۳۲۵	۰/۳۲۵	۰/۱۱۷	۰/۱۱۷	۰/۱۱۷
λ_{17}	۰/۴۴۴	۰/۴۴۳	۰/۴۴۲	۰/۱۱۶	۰/۱۱۶	۰/۱۱۶
λ_{18}	۰/۲۳۴	۰/۲۳۵	۰/۲۳۷	۰/۱۳۸	۰/۱۳۸	۰/۱۳۸
λ_{19}	۰/۸۰۹	۰/۸۰۸	۰/۸۰۷	۰/۰۹۷	۰/۰۹۷	۰/۰۹۷
λ_{20}	۰/۳۲۷	۰/۳۲۹	۰/۳۳۰	۰/۱۲۴	۰/۱۲۴	۰/۱۲۴
λ_{21}	۰/۸۱۲	۰/۸۱۱	۰/۸۱۰	۰/۰۷۵	۰/۰۷۵	۰/۰۷۵

k	$\hat{\theta}_k$		
	۰/۵	۰/۶	۰/۷
ψ_{11}	۰/۵۵۹	۰/۵۶۰	۰/۵۶۰
ψ_{22}	۰/۴۵۹	۰/۴۵۱	۰/۴۵۰
ψ_{33}	۰/۳۱۶	۰/۳۱۷	۰/۳۱۸
ψ_{44}	۰/۹۰۵	۰/۹۰۴	۰/۹۰۳
ψ_{55}	۰/۴۹۴	۰/۴۹۲	۰/۴۹۰
ψ_{66}	۰/۶۹۷	۰/۶۹۵	۰/۶۹۴
ψ_{77}	۰/۵۷۴	۰/۵۷۳	۰/۵۷۲
ψ_{88}	۰/۵۱۸	۰/۵۱۸	۰/۵۱۷
ψ_{99}	۰/۵۸۹	۰/۵۸۷	۰/۵۸۵
$\psi_{10,10}$	۰/۳۰۲	۰/۳۰۰	۰/۲۹۸
$\psi_{11,11}$	۰/۶۹۹	۰/۶۹۸	۰/۶۹۸
$\psi_{12,12}$	۰/۶۴۶	۰/۶۴۵	۰/۶۴۴
$\psi_{13,13}$	۰/۵۶۷	۰/۵۶۸	۰/۵۶۹
$\psi_{14,14}$	۰/۳۴۳	۰/۳۴۶	۰/۳۴۸
$\psi_{15,15}$	۰/۵۷۶	۰/۵۷۹	۰/۵۸۱
$\psi_{16,16}$	۰/۸۹۵	۰/۸۹۵	۰/۸۹۴
$\psi_{17,17}$	۰/۸۰۳	۰/۸۰۴	۰/۸۰۵
$\psi_{18,18}$	۰/۹۴۵	۰/۹۴۵	۰/۹۴۴
$\psi_{19,19}$	۰/۳۴۶	۰/۳۴۸	۰/۳۵۰
$\psi_{20,20}$	۰/۸۹۳	۰/۸۹۲	۰/۸۹۱
$\psi_{21,21}$	۰/۳۴۰	۰/۳۴۲	۰/۳۴۳

جدول ۲.۲: آماره‌هایی برای ارزیابی مدل

k	T_{RML_k}	p	T_{AML_k}	p	\hat{m}_2	T_{RRLS_k}	p	T_{RLS_k}	p
۰/۵	۴۸۰/۷۰	۰/۰۰۰	۱۱۰/۴۹	۰/۰۰۰	۴۳/۴۴	۲۹۲/۶۸	۰/۰۰۰	۶۷/۲۸	۰/۰۱۲
۰/۶	۳۸۱/۹۴	۰/۰۰۰	۸۷/۷۹	۰/۰۰۰	۴۳/۴۴	۲۹۳/۸۸	۰/۰۰۰	۶۷/۵۵	۰/۰۱۱
۰/۷	۳۵۰/۷۰	۰/۰۰۰	۸۰/۶۰	۰/۰۰۱	۴۳/۴۴	۲۹۴/۹۰	۰/۰۰۰	۶۷/۷۷	۰/۰۱۱

به‌عنوان نکته‌ی آخر، روش توسعه یافته می‌تواند به راحتی در نرم‌افزارهایی که در حال حاضر دارای این گزینه از مدل‌های ماتریس همبستگی چندحالتی، چندرشته‌ای و گشتاور- حاصلضربی اجرا شود. روش ML استوار توسط محصول R19 اجرا می‌شود. با قرار دادن R_k به جای R ، تنها نیاز است که روی قطر اصلی برازش مدل، $k+1$ را به جای ۱ در فرآیند تکرار جای‌گذاری کرد در نتیجه آماره‌های T_{MLS} ، T_{AML} ، T_{RML} و T_{ML} به‌طور خودکار به ترتیب با آماره‌های T_{MLS_k} ، T_{AML_k} ، T_{RML_k} و T_{ML_k} جایگذاری می‌شوند. با دانش امروزی انسان هیچ نرم‌افزاری قادر به تولید دقیق T_{ARLS_k} و T_{RRLS_k} نیست، با این حال با T_{RRLS_k} ، \hat{m}_1 و \hat{m}_2 این دو آماره به راحتی می‌توانند تولید (محاسبه) شوند.

اگرچه در این جا (به معنای واقعی کلمه)، R_k یک ماتریس کواریانس است، به‌جز $\hat{\theta}_k$ غیر استاندارد، زمانی که $diag(R_k) = \Sigma(\hat{\theta}_k)$ رخ می‌دهد، رفتار R_k به‌عنوان یک ماتریس کواریانس، تجزیه و تحلیل درستی را تولید نخواهد کرد.

فصل ۳

مدل معادلات ساختاری جریمه شده

۱.۳ مقدمه

در خصوص ساده‌سازی ساختار مدل تحقیقات زیادی صورت گرفته است. ساختار ساده توسط تارتستون^۱ (۱۹۳۵) مطرح شد. پیچیدگی متغیر را براون^۲ (۲۰۰۱) مورد بررسی قرار داد. در تحلیل مولفه‌های اصلی به کاربران بارهای عاملی صفر را ژو و همکاران^۳ (۲۰۰۶) پیشنهاد کردند. جاکوبوچی و همکاران^۴ (۲۰۱۶) مدل معادلات ساختاری جریمه‌شده را ارائه کردند و در ادامه جاکوبوچی^۵ (۲۰۱۷) بسته regsem را در نرم‌افزار R برای تحلیل این‌گونه مدل‌ها معرفی کرد. هدف در این قسمت برآوردی دقیق و کارا از مدلی ساده است که به کاربر این اجازه را می‌دهد به راحتی مدل واقعی را تفسیر کند. برای این منظور در این فصل مدل معادلات ساختاری جریمه شده جاکوبوچی (۲۰۱۷) را مورد تحلیل و بررسی قرار می‌دهیم.

ایده اصلی، همانند همه‌ی مدل‌های جریمه شده، این است که برای ساده‌سازی مدل و حذف اثرات بی‌معنی، یک جمله جریمه به تابع هدف اضافه شود. با توجه به این که در فصل ۲، در برآورد پارامترهای SEM از تابع هدف بر پایه‌ی درست‌نمایی (FML) استفاده کردیم، در

^۱Thurstone

^۲Brown

^۳Zou et al

^۴Jacobucci et al

^۵Jacobucci

این فصل SEM را از دیدگاه درست‌نمایی جریمه تحلیل می‌کنیم. لازم به ذکر است که در فصل ۳ اگرچه برآورد ریج در SEM مطرح شد ولی دیدگاهی متفاوت از مدل جریمه شده بود. در این فصل برآورد ریج را از دیدگاه SEM جریمه شده مورد بررسی قرار داده و برای کلیت، برخی توابع جریمه معروف دیگر را نیز در بررسی‌های خود در نظر می‌گیریم.

۲.۳ درست‌نمایی جریمه شده

در برآورد پارامترهای SEM تابع هدف بر پایه‌ی درست‌نمایی به صورت

$$F_{ML}(\theta) = \log(|\Sigma(\theta)|) + tr(S\Sigma^{-1}(\theta)) - \log(|S|) - p \quad (1.3)$$

را در نظر بگیرید که در آن S ماتریس کواریانس نمونه و $\Sigma(\theta)$ ماتریس کواریانس مدل به صورت

$$\Sigma(\theta) = \Lambda\Phi\Lambda^T + \Psi$$

است. در این جا p تعداد پارامترهای مدل می‌باشد. همان طور که در فصل ۲ اشاره شد برآوردگر ML پارامتر $\theta = (\theta_1, \dots, \theta_p)^T$ عبارتست از

$$\hat{\theta} = \underset{\theta}{\operatorname{argmin}} F_{ML}(\theta)$$

رابطه (۱.۳) تابع هدف جریمه شده به صورت $\hat{\theta} = \underset{\theta}{\operatorname{argmin}} F_{ML}(\theta)$ بر پایه‌ی،

$$F_{regsem}(\theta) = F_{ML}(\theta) + \lambda P(\theta)$$

را در نظر بگیرید که در آن λ پارامتر تنظیم کننده است که مقداری بین صفر تا بی‌نهایت اختیار می‌کند و $P(\theta)$ نشان دهنده یک تابع جریمه کلی برای پارامتر θ می‌باشد. به ازای $\lambda = 0$ ، $F_{regsem}(\theta) = F_{ML}(\theta)$ و وقتی $\lambda \rightarrow \infty$ ، همه‌ی پارامترها صفر برآورد می‌شوند (به سمت صفر منقبض می‌شوند). در این صورت برآوردگر جریمه شده θ عبارتست از

$$\hat{\theta}_{regsem} = \underset{\theta}{\operatorname{argmin}} F_{regsem}(\theta) \quad (2.3)$$

لازم به ذکر است، به ازای هیچ تابع جریمه‌ای، صورت بسته‌ای برای جواب رابطه (۲.۳) وجود ندارد. لذا جاکوبوچی (۲۰۱۷) برای تعیین جواب مساله بهینه‌سازی مورد نظر بسته regsem در نرم‌افزار R طراحی کرد که با استفاده از آن بتوان جواب رابطه (۲.۳) را به ازای برخی توابع جریمه معروف به دست آورد. در ادامه این فصل روش کار بسته regsem را شرح داده و در قالب مثالی خروجی آن را تحلیل می‌کنیم. اما قبل از آن لازم است به‌طور بسیار مختصری توابع جریمه معروف که توسط این بسته پوشش داده می‌شوند را معرفی کنیم.

۱.۲.۳ انواع توابع جریمه

معروف‌ترین روش‌های جریمه در رگرسیون عبارتند از روش رگرسیون ریج (هورل و کنارد)^۶، (۱۹۷۰) و لاسو (تیبشیرانی^۷، ۱۹۹۶). بدیهی است برخی از روش‌های جریمه را می‌توان به صورت تعمیمی از این دو روش در نظر گرفت.

ریج

با توجه به تابع هدف SEM برآورد ریج به صورت زیر تعریف می‌شود

$$\hat{\theta}^{ridge} = \underset{\theta}{\operatorname{argmin}} \left\{ F_{ML}(\theta) + \lambda \sum_{j=1}^p \theta_j^2 \right\}, \quad (۳.۳)$$

که در آن λ یک پارامتر ریج است که مقدار انقباض را کنترل می‌کند. لازم به یادآوری است چنانچه برآورد ریج را در مدل رگرسیون چندگانه (۱.۳) محاسبه می‌کردیم، صورت بسته‌ای برای جواب داشتیم. همان‌طور که در فصل ۲ نشان دادیم

$$\hat{\beta}^{ridge} = \underset{\theta}{\operatorname{argmin}} \left\{ (\mathbf{y} - \mathbf{x}\beta)^T (\mathbf{y} - \mathbf{x}\beta) + \lambda \sum_{j=1}^p \beta_j^2 \right\} = (\mathbf{x}^T \mathbf{x} + k \mathbf{I}_p)^{-1} \mathbf{x}^T \mathbf{y} \quad (۴.۳)$$

لاسو

برآورد لاسو به این صورت تعریف می‌شود

$$\hat{\theta}^{lasso} = \underset{\theta}{\operatorname{argmin}} \left\{ F_{ML}(\theta) + \lambda \sum_{j=1}^p |\theta_j| \right\}, \quad (۵.۳)$$

برای آگاهی بیشتر در خصوص برآورد لاسو در مدل‌های خطی به آرست^۸ (۱۳۹۵) مراجعه کنید.

الاستیک نت

جریمه الاستیک‌نت توسط ژو و هستی^۹ (۲۰۰۵) برای اولین بار ارائه شد. این جریمه هر دو روش ریج و لاسو را شامل می‌شود. این روش از طریق اضافه کردن پارامتر اضافی α ، بین روش ریج و لاسو یک تعامل (سازش) ایجاد می‌کند. جریمه الاستیک‌نت به صورت زیر نوشته می‌شود

$$P^{Enet}(\theta_j) = (1 - \alpha) \|\theta_j\|_2 + \alpha \|\theta_j\|_1$$

^۶ Hoerl and Kennard

^۷ Tibshirani

^۸ Arast

^۹ Zou and Hastie

که در آن $\|\theta_j\|_1 = |\theta_j|$ و $\|\theta_j\|_1 = \theta_j^2$ است بنابراین برآورد الاستیکنت عبارتست از

$$\hat{\theta}^{Enet} = \underset{\theta}{\operatorname{argmin}} \left\{ F_{ML}(\theta) + \lambda \sum_{j=1}^p p^{Enet}(\theta_j) \right\}, \quad (6.3)$$

زمانی که $\alpha = 0$ باشد، همان جریمه ریج و برعکس زمانی که $\alpha = 1$ باشد، همان جریمه لاسو حاصل می‌شود. برتری عمومی الاستیکنت به لاسو به این است که برای داده‌هایی که وابستگی زیادی دارند، به دلیل وجود جمله جریمه ریج، کارایی بهتری دارد.

لاسوی تطبیقی

در استفاده از جریمه لاسو، زمانی که مقیاس متغیرها به‌طور چشمگیری متفاوت باشد مشکلاتی به‌وجود می‌آید. تنها با اضافه کردن یک مقدار λ می‌توان ارزیابی قابل ملاحظه‌ای را به برآوردهای حاصل اضافه کرد (فن و لی ۲۰۰۱). برآورد لاسوی تطبیقی به‌صورت زیر تعریف می‌شود

$$\hat{\theta}^{ALasso} = \underset{\theta}{\operatorname{argmin}} \left\{ F_{ML}(\theta) + \lambda \sum_{j=1}^p \hat{W}_j |\theta_j| \right\},$$

که در آن

$$\hat{W}_j = \frac{1}{|\hat{\theta}_j|^r}, \quad r > 0$$

و $\hat{\theta}_j$ مولفه j ام برآورد ML پارامتر θ است. یکی از مزیت‌های برآورد لاسوی تطبیقی این است که برخلاف لاسو خاصیت خوب توزیع نرمال مجانبی را دارا می‌باشد.

SCAD

فن و لی (۲۰۰۱) برای اولین بار، بر خلاف توابع جریمه‌ای که در بالا به آن‌ها اشاره شد، تابع جریمه‌ای مقعر به نام اسکد را معرفی کردند. تابع جریمه اسکد شامل ۲ پارامتر تنظیم است که به‌صورت زیر می‌باشد

$$p^{SCAD}(\theta_j) = \begin{cases} \lambda |\theta_j|, & |\theta_j| \leq \lambda \\ -\frac{\theta_j^2 - 2a\lambda|\theta_j| + \lambda^2}{2(a-1)}, & \lambda < |\theta_j| \leq a\lambda \\ \frac{(a+1)\lambda^2}{2}, & |\theta_j| > a\lambda \end{cases}$$

که در عمل مقدار بهینه $a = 3/7$ مورد استفاده قرار می‌گیرد. جریمه اسکد به‌طور پیوسته روی $(-\infty, 0) \cup (0, \infty)$ مشتق‌پذیر است. جریمه اسکد می‌تواند تعداد زیادی از پارامترها را صفر کند

و ضرایب تقریبا نارایب را برای غیر صفرها ارائه دهد. در این حالت برآوردگر اسکد در SEM عبارتست از

$$\hat{\theta}^{SCAD} = \underset{\theta}{\operatorname{argmin}} \left\{ F_{ML}(\theta) + \lambda \sum_{j=1}^p p^{SCAD}(\theta_j) \right\},$$

MCP

ژانگ^{۱۱} (۲۰۱۰) جریمه مقعر مینیماکس (MCP) را به صورت زیر معرفی کرد

$$p^{MCP}(\theta_j) = \lambda \left(|\theta_j| - \frac{\theta_j^2}{2\lambda r} \right) \mathbf{I}(|\theta_j| < \lambda r) + \frac{\lambda^2 r}{2} \mathbf{I}(|\theta_j| \geq \lambda r)$$

که در آن $\mathbf{I}(A)$ نشان دهنده تابع نشانگر مجموعه A می باشد. در عمل مقدار بهینه $r = 3/7$ مورد استفاده قرار می گیرد.

۳.۳ بسته regsem

جاکوبوچی بسته regsem را در نرم افزار R برای اجرای SEM جریمه شده معرفی کرد. این بسته از روش RAM (مک آردل^{۱۲}، ۲۰۰۵) برای تجزیه و تحلیل استفاده می کند. پارامترهای SEM به سه ماتریس تفکیک می شوند: فیلتر (F)، نامتقارن (A)؛ مسیر مستقیم؛ به عنوان مثال بارهای عاملی یا رگرسیون)، و متقارن (S)؛ مسیر غیر مستقیم؛ به عنوان مثال واریانس ها یا کواریانس ها). برای اطلاع از جزئیات بیشتر در مورد نماد RAM و کاربرد آن در Regsem به جاکوبوچی و همکاران^{۱۳} (۲۰۱۶) مراجعه کنید.

نحوه استفاده از regsem براساس بسته lavaan (روسل^{۱۴}، ۲۰۱۲)، برای مدل های معادلات ساختاری می باشد. lavaan یک برنامه نرم افزاری برای مدل سازی معادلات ساختاری کلی (عمومی) است که می تواند یک بعد گسترده ای از مدل ها با روش های برآورد متنوع مختلفی را برازش دهد.

برای استفاده از regsem کاربر ابتدا مدل را در lavaan برازش می دهد. توجه داشته باشید به ویژه در مواردی که تعداد متغیرها بزرگتر از حجم نمونه است. در بسته lavaan مدل نیاز به همگرایی ندارد، چه رسد به اجرا. در این مورد، از آرگومان $do.fit = FALSE$ در بسته lavaan می توان استفاده کرد.

^{۱۱}Zhang

^{۱۲}McArdle

^{۱۳}Jacobucci

^{۱۴}Rossee

به عنوان مثالی معیار (متعارف)، در زیر کدی برای یک مدل تحلیل عاملی تائیدی با یک عامل پنهان و نه شاخص از مجموعه داده هولزینگر و اسوین فورد^{۱۵} (۱۹۳۹) ارائه می‌کنیم.

```
library(lavaan)
mod <- '
f1 =~NA*x1 + x2 + x3 + x4 + x5 + x6 + x7 + x8 + x9
f1~~1*f1
'
out <- cfa(mod,HolzingerSwineford1939,meanstructure=TRUE)
```

پس از این که مدل در lavaan اجرا شد، از lavaan() یا هر یک از توابع بسته‌بندی برای برازش یک مدل (به عنوان مثال، sem()، cfa()، یا growth())، موردی که از طریق بسته regsem() برای تفکیک مدل به نماد RAM استفاده می‌شود و با استفاده از یکی از سه تابع regsem()، multi_optim()، یا cv_regsem() اجرا می‌شود. تابع regsem() مدل را با یک مقدار جریمه اجرا می‌کند، در حالی که تابع multi_optim() همان کار را انجام می‌دهد با این تفاوت که اجازه می‌دهد برای شروع از مقادیر تصادفی استفاده شود. با این حال، تابع اصلی cv_regsem() نه تنها مدل را اجرا می‌کند، که در سرتاسر یک بردار از مقادیر جریمه مختلف برای اجرا استفاده می‌کند.

حال از برآورد Ridge استفاده می‌کنیم (type="ridge"). به عنوان مثال در مدل تک عاملی بالا، هر یک از بارهای عاملی می‌توانند با جریمه‌های Ridge آزمون شوند برای تعیین این که هر یک از شاخص‌ها بخشی ضروری (لازم) از عامل پنهان است داریم

```
library(regsem)
extractMatrices(out) ["A"]
out.reg <- cv_regsem(out, type="ridge",optMethod="coord_desc",
pars_pen=c(1:9), jump=.05)
```

در این جا، تابع extractMatrices()، به کاربر اجازه می‌دهد بررسی کند که چگونه مدل lavaan ماتریس RAM را تفسیر می‌کند. علاوه بر این با نگاه کردن به ماتریس A، تعداد پارامتر مربوط به بارهای عاملی مورد علاقه برای تنظیم را می‌توان شناسایی کرد. برای این مدل، بارهای عاملی نشان‌دهنده‌ی تعداد پارامترهای ۱ تا ۹ می‌باشند، که به‌طور مستقیم از آرگومان pars_pen تابع cv_regsem() (اگر pars_pen = NULL باشد آن‌گاه همه‌ی اثرات مستقیم، خارج از برش (عرض از مبدا)، جریمه می‌شوند.) استفاده می‌کنیم. علاوه بر این، یک سلسله دلایل قابل قبول از این که چگونه مقادیر زیادی از جریمه را می‌خواهیم از طریق n.lambda

^{۱۵}Holzinger and Swineford

(15 = آزمون کنیم، چگونه مقادیر جریمه برای هر مدل باید افزایش یابد (jump=0.05)، و در نهایت از برآورد Ridge استفاده شده است (type="ridge"). out.reg دو بخش را شامل می‌شود، out.reg[[1]] برآورد پارامتر برای هر یک از ۱۰۰ مدل است

```
head(round(out.reg[[1]],2),5)
f1 -> x1 f1 -> x2 f1 -> x3 f1 -> x4 f1 -> x5 f1 -> x6 f1 -> x7 f1 -> x8
[1,] 0.51 0.26 0.25 0.98 1.08 0.92 0.20 0.20
[2,] 0.45 0.21 0.20 0.92 1.01 0.85 0.15 0.16
[3,] 0.36 0.18 0.18 0.78 0.84 0.73 0.14 0.15
[4,] 0.33 0.16 0.16 0.73 0.78 0.68 0.13 0.13
[5,] 0.30 0.15 0.14 0.69 0.73 0.65 0.12 0.12
f1 -> x9 1 -> x1 1 -> x2 1 -> x3 1 -> x4 1 -> x5 1 -> x6 1 -> x7 1 -> x8
[1,] 0.31 4.94 6.09 2.25 3.06 4.34 2.19 4.19 5.53
[2,] 0.26 4.94 6.09 2.25 3.06 4.34 2.19 4.19 5.53
[3,] 0.22 4.94 6.09 2.25 3.06 4.34 2.19 4.19 5.53
[4,] 0.20 4.94 6.09 2.25 3.06 4.34 2.19 4.19 5.53
[5,] 0.19 4.94 6.09 2.25 3.06 4.34 2.19 4.19 5.53
1 -> x9 x1 ~~ x1 x2 ~~ x2 x3 ~~ x3 x4 ~~ x4 x5 ~~ x5 x6 ~~ x6 x7 ~~ x7
[1,] 5.37 1.10 1.31 1.21 0.38 0.49 0.36 1.15
[2,] 5.37 1.10 1.31 1.21 0.38 0.48 0.36 1.15
[3,] 5.37 1.11 1.32 1.22 0.39 0.51 0.36 1.15
[4,] 5.37 1.12 1.32 1.22 0.40 0.52 0.37 1.15
[5,] 5.37 1.12 1.32 1.22 0.41 0.54 0.37 1.15
x8 ~~ x8 x9 ~~ x9
[1,] 0.98 0.92
[2,] 0.98 0.92
[3,] 0.98 0.92
[4,] 0.98 0.93
[5,] 0.99 0.93
```

در حالی که out.reg[[2]]، اطلاعات مربوط به برازش هر مدل را شامل می‌شود

```
round(out.reg[[2]],2)
lambda conv rmsea BIC
```

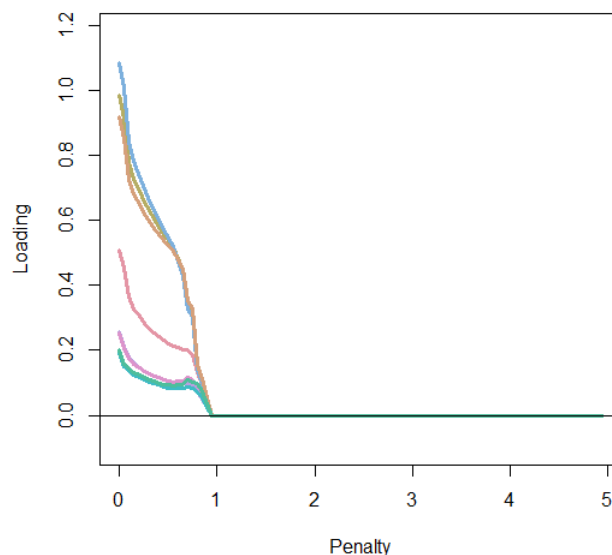
[1,]	0.00	0	0.19	7856.54
[2,]	0.05	0	0.19	7860.71
[3,]	0.10	0	0.20	7888.42
[4,]	0.15	0	0.20	7909.83
[5,]	0.20	0	0.21	7931.42
[6,]	0.25	0	0.22	7953.05
[7,]	0.30	0	0.22	7974.66
[8,]	0.35	0	0.23	7996.33
[9,]	0.40	0	0.23	8018.16
[10,]	0.45	0	0.24	8040.37
[11,]	0.50	0	0.25	8063.06
[12,]	0.55	0	0.25	8086.51
[13,]	0.60	0	0.26	8122.04
[14,]	0.65	0	0.27	8160.58
[15,]	0.70	0	0.29	8252.61
[16,]	0.75	0	0.30	8280.01
[17,]	0.80	0	0.32	8398.64
[18,]	0.85	0	0.33	8428.68
[19,]	0.90	0	0.33	8455.64
[20,]	0.95	0	0.33	8463.13
[21,]	1.00	0	0.33	8463.13
[22,]	1.05	0	0.33	8463.13
[23,]	1.10	0	0.33	8463.13
[24,]	1.15	0	0.33	8463.13
[25,]	1.20	0	0.33	8463.13
[26,]	1.25	0	0.33	8463.13
[27,]	1.30	0	0.33	8463.13
[28,]	1.35	0	0.33	8463.13
[29,]	1.40	0	0.33	8463.13
[30,]	1.45	0	0.33	8463.13
[31,]	1.50	0	0.33	8463.13
[32,]	1.55	0	0.33	8463.13
[33,]	1.60	0	0.33	8463.13
[34,]	1.65	0	0.33	8463.13

[35,] 1.70 0 0.33 8463.13
[36,] 1.75 0 0.33 8463.13
[37,] 1.80 0 0.33 8463.13
[38,] 1.85 0 0.33 8463.13
[39,] 1.90 0 0.33 8463.13
[40,] 1.95 0 0.33 8463.13
[41,] 2.00 0 0.33 8463.13
[42,] 2.05 0 0.33 8463.13
[43,] 2.10 0 0.33 8463.13
[44,] 2.15 0 0.33 8463.13
[45,] 2.20 0 0.33 8463.13
[46,] 2.25 0 0.33 8463.13
[47,] 2.30 0 0.33 8463.13
[48,] 2.35 0 0.33 8463.13
[49,] 2.40 0 0.33 8463.13
[50,] 2.45 0 0.33 8463.13
[51,] 2.50 0 0.33 8463.13
[52,] 2.55 0 0.33 8463.13
[53,] 2.60 0 0.33 8463.13
[54,] 2.65 0 0.33 8463.13
[55,] 2.70 0 0.33 8463.13
[56,] 2.75 0 0.33 8463.13
[57,] 2.80 0 0.33 8463.13
[58,] 2.85 0 0.33 8463.13
[59,] 2.90 0 0.33 8463.13
[60,] 2.95 0 0.33 8463.13
[61,] 3.00 0 0.33 8463.13
[62,] 3.05 0 0.33 8463.13
[63,] 3.10 0 0.33 8463.13
[64,] 3.15 0 0.33 8463.13
[65,] 3.20 0 0.33 8463.13
[66,] 3.25 0 0.33 8463.13
[67,] 3.30 0 0.33 8463.13
[68,] 3.35 0 0.33 8463.13

[69,] 3.40 0 0.33 8463.13
[70,] 3.45 0 0.33 8463.13
[71,] 3.50 0 0.33 8463.13
[72,] 3.55 0 0.33 8463.13
[73,] 3.60 0 0.33 8463.13
[74,] 3.65 0 0.33 8463.13
[75,] 3.70 0 0.33 8463.13
[76,] 3.75 0 0.33 8463.13
[77,] 3.80 0 0.33 8463.13
[78,] 3.85 0 0.33 8463.13
[79,] 3.90 0 0.33 8463.13
[80,] 3.95 0 0.33 8463.13
[81,] 4.00 0 0.33 8463.13
[82,] 4.05 0 0.33 8463.13
[83,] 4.10 0 0.33 8463.13
[84,] 4.15 0 0.33 8463.13
[85,] 4.20 0 0.33 8463.13
[86,] 4.25 0 0.33 8463.13
[87,] 4.30 0 0.33 8463.13
[88,] 4.35 0 0.33 8463.13
[89,] 4.40 0 0.33 8463.13
[90,] 4.45 0 0.33 8463.13
[91,] 4.50 0 0.33 8463.13
[92,] 4.55 0 0.33 8463.13
[93,] 4.60 0 0.33 8463.13
[94,] 4.65 0 0.33 8463.13
[95,] 4.70 0 0.33 8463.13
[96,] 4.75 0 0.33 8463.13
[97,] 4.80 0 0.33 8463.13
[98,] 4.85 0 0.33 8463.13
[99,] 4.90 0 0.33 8463.26
[100,] 4.95 0 0.33 8463.23

در بررسی جریمه (lambda)، چنانچه (conv=0) باشد، در این صورت مدل همگراست. به طور پیش فرض دو شاخص برای خروجی مناسب هستند، ریشه‌ی میانگین مربعات خطای تقریبی (RMSEA؛ استیگر و لیند^{۱۶} ۱۹۸۰) و معیار اطلاع بیز (BIC اسچوارز^{۱۷}، ۱۹۷۸). هر دو شاخص RMSEA و BIC درجه آزادی مدل را محاسبه می‌کنند، نکته‌ی مهم برای انتخاب مدل در حضور جریمه ریج (و دیگر جریمه‌ها که مجموعه پارامترها را به سمت صفر می‌برد) است. ژو و همکاران^{۱۸} (۲۰۰۷) ثابت کردند که تعداد پارامترهای غیرصفر، برآوردی ناریب از درجه آزادی برای رگرسیون است. با افزایش مقدار جریمه، مجموعه پارامترهایی انتخاب می‌شود که به صفر میل می‌کنند، بنابراین درجه آزادی را برای شاخص‌های برازشی که شامل محاسبه درجه آزادی می‌شوند، افزایش می‌دهیم. به این معنا که درست‌نمایی مدل ممکن است بد و بدتر شود (افزایش یابد)، هر دو شاخص RMSEA و BIC می‌توانند بهبود یابند (کاهش یابند). به جای بررسی ماتریس خروجی برآورد پارامترها [out.reg[[1]]، کاربران گزینه‌ای برای رسم خط سیر (گذرگاه) هر یک از پارامترهای جریمه دارند. برای این کار از دستور (۹ : ۱ = pars) `plot_cv(out.reg, pars = 1 : 9)` استفاده شده است، و در نتیجه داریم

`plot_cv(out.reg, pars=1:9)`



شکل ۱.۳: نمودار اثر ریج داده‌های بخش ۳.۳

پس از این که یک مدل نهایی (جریمه) انتخاب شد، کاربران گزینه‌ای دارند که فقط از خروجی `cv_regsem` استفاده می‌کند، یا مدل نهایی را برای رسیدن به اطلاعات اضافی می‌توان

^{۱۶}Steiger and Lind

^{۱۷}Schwarz

^{۱۸}Zou et al

دوباره با دستورات `regsem()` و `multi_optim()` اجرا کرد. در مدل بالا بهترین جریمه بر اساس شاخص BIC مقدار $\lambda = 0$ می باشد. با این حال، برای اثبات این موضوع می توانیم مقدار جریمه $0/2$ را انتخاب کنیم. حال مدل را بار دیگر با دستور `regsem()` اجرا می کنیم.

```
mod.out <- regsem(out, type="ridge", pars_pen=c(1:9),lambda=0.2)
#summary(mod.out)
```

توجه داشته باشید که آرگومان های استفاده شده در بالا، مربوط به همان آرگومان ها در `multi_optim()` می باشد. با این حال، تابع `multi_optim()` آرگومان های اضافی اختیاری مربوط به تعداد یا شروع با اعداد تصادفی برای آزمون را دارا می باشد. شاخص های برازش اضافی را می توان از طریق تابع `fit_indices` به دست آورد.

```
fit_indices(mod.out)
$Data_Type
[1] "Train"

$fits
Fmin varFit p chisq p.chisq
0.54991 0.00000 9.00000 331.04623 0.00000
nfac df npar N baseline.chisq
1.00000 27.00000 27.00000 301.00000 918.85159
baseline.df logl ncp rmsea rmsea.lower
36.00000 -3860.61528 1.01349 0.19374 0.17508
rmsea.upper rmsea.pval CFI TLI BIC
NA 0.00000 0.65561 0.54081 7875.32253
AIC CAIC EBIC.5 EBIC.25
7775.23056 7875.41209 7934.64760 7904.98507
```

با تغییر پیش فرض آرگومان ("`RMSEA`" , "`BIC`") از طریق تابع `CV_regsem` می توان به همین اندازه های برازش دست یافت (دسترسی پیدا کرد). در نهایت، به جای ارزیابی این شاخص های برازش در نمونه های مشابه که مدل روی آن اجرا شد، از یک مجموعه داده گسترده می توان استفاده کرد.

این کار را می توان به دو روش انجام داد: هم با `cv_regsem(..., fit.ret = "test")` و هم با `fit_indices(model, CV = TRUE, CovMat =)` و نام ماتریس کواریانس گسترده را تعیین می کند.

حال از برآورد lasso استفاده می‌کنیم (`type="lasso"`). به عنوان مثال در مدل تک عاملی بالا، هر یک از بارهای عاملی می‌توانند با جریمه‌های lasso آزمون شوند برای تعیین این‌که هر یک از شاخص‌ها بخشی ضروری (لازم) از عامل پنهان است داریم

```
library(regsem)
extractMatrices(out) ["A"]
out.reg <- cv_regsem(out, type="lasso", pars_pen=c(1:9), n.lambda=15, jump=.05)
```

یک سلسله دلایل قابل قبول از این‌که چگونه مقادیر زیادی از جریمه را می‌خواهیم از طریق ($n.lambda = 15$) آزمون کنیم، چگونه مقادیر جریمه برای هر مدل باید افزایش یابد ($jump = 0.05$)، و در نهایت از برآورد lasso استفاده شده است (`type="lasso"`). `out.reg` دو بخش را شامل می‌شود، `out.reg[[1]]` برآورد پارامتر برای هر یک از ۱۵ مدل است،

```
head(round(out.reg[[1]], 2), 5)
f1 -> x1 f1 -> x2 f1 -> x3 f1 -> x4 f1 -> x5 f1 -> x6 f1 -> x7 f1 -> x8
[1,] 0.51 0.26 0.25 0.98 1.08 0.92 0.20 0.20
[2,] 0.41 0.17 0.17 0.89 0.98 0.83 0.12 0.13
[3,] 0.33 0.10 0.10 0.83 0.91 0.77 0.06 0.08
[4,] 0.27 0.04 0.04 0.78 0.87 0.72 0.01 0.03
[5,] 0.21 0.00 0.00 0.75 0.83 0.69 0.00 0.00
f1 -> x9 1 -> x1 1 -> x2 1 -> x3 1 -> x4 1 -> x5 1 -> x6 1 -> x7 1 -> x8
[1,] 0.31 4.94 6.09 2.25 3.06 4.34 2.19 4.19 5.53
[2,] 0.23 4.94 6.09 2.25 3.06 4.34 2.19 4.19 5.53
[3,] 0.17 4.94 6.09 2.25 3.06 4.34 2.19 4.19 5.53
[4,] 0.12 4.94 6.09 2.25 3.06 4.34 2.19 4.19 5.53
[5,] 0.07 4.94 6.09 2.25 3.06 4.34 2.19 4.19 5.53
1 -> x9 x1 ~ x1 x2 ~ x2 x3 ~ x3 x4 ~ x4 x5 ~ x5 x6 ~ x6 x7 ~ x7
[1,] 5.37 1.10 1.31 1.21 0.38 0.49 0.36 1.15
[2,] 5.37 1.11 1.32 1.22 0.38 0.48 0.36 1.15
[3,] 5.37 1.13 1.33 1.23 0.38 0.47 0.36 1.16
[4,] 5.37 1.15 1.35 1.25 0.38 0.46 0.37 1.17
[5,] 5.37 1.18 1.36 1.26 0.39 0.46 0.38 1.18
x8 ~ x8 x9 ~ x9
[1,] 0.98 0.92
[2,] 0.99 0.93
```

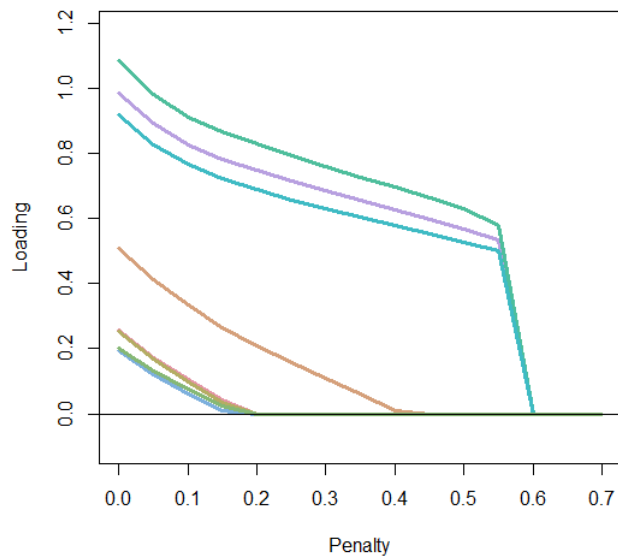
```
[3,] 1.00 0.94
[4,] 1.01 0.96
[5,] 1.02 0.97
```

در حالی که `out.reg[[2]]` ، اطلاعات مربوط به برازش هر مدل را شامل می‌شود

```
round(out.reg[[2]],2)
lambda conv rmsea BIC
[1,] 0.00 0 0.19 7856.54
[2,] 0.05 0 0.19 7866.75
[3,] 0.10 0 0.20 7891.95
[4,] 0.15 0 0.21 7927.21
[5,] 0.20 0 0.20 7937.51
[6,] 0.25 0 0.21 7963.18
[7,] 0.30 0 0.21 7985.82
[8,] 0.35 0 0.22 8012.03
[9,] 0.40 0 0.23 8042.40
[10,] 0.45 0 0.23 8060.78
[11,] 0.50 0 0.23 8086.17
[12,] 0.55 0 0.24 8120.85
[13,] 0.60 0 0.29 8411.90
[14,] 0.65 0 0.29 8411.89
[15,] 0.70 0 0.29 8411.89
```

به‌جای بررسی ماتریس خروجی برآورد پارامترها `out.reg[[1]]`، کاربران گزینه‌ای برای رسم خط سیر (گذرگاه) هر یک از پارامترهای جریمه دارند. برای این کار از دستور `plot_cv(out.reg, pars = ۱ : ۹)` استفاده شده است، و در نتیجه داریم

```
plot_cv(out.reg, pars=1:9)
```

شکل ۲.۳: نمودار اثر لاسو داده‌های بخش ۳.۳

پس از این که یک مدل نهایی (جریمه) انتخاب شد، کاربران گزینه‌ای دارند که فقط از خروجی *cv_regsem* استفاده می‌کند، یا مدل نهایی را برای رسیدن به اطلاعات اضافی می‌توان دوباره با دستورات *regsem()* و *multi_optim()* اجرا کرد. در مدل بالا بهترین جریمه بر اساس شاخص BIC مقدار $\lambda = 0$ می‌باشد. با این حال، برای اثبات این موضوع می‌توانیم مقدار جریمه 0.2 را انتخاب کنیم. حال مدل را بار دیگر با دستور *regsem()* اجرا می‌کنیم.

```
mod.out <- regsem(out, type="lasso", pars_pen=c(1:9),lambda=0.2)
#summary(mod.out)
```

شاخص‌های برازش اضافی را می‌توان از طریق تابع *fit_indices* به‌دست آورد.

```
fit_indices(mod.out)
$Data_Type
[1] "Train"

$fits
Fmin varFit p chisq p.chisq
0.69087 0.00000 9.00000 415.90286 0.00000
nfac df npar N baseline.chisq
```

```
1.00000 31.00000 23.00000 301.00000 918.85159
baseline.df logl ncp rmsea rmsea.lower
36.00000 -3903.04360 1.28301 0.20344 0.18598
rmsea.upper rmsea.pval CFI TLI BIC
NA 0.00000 0.56402 0.49370 7937.35073
AIC CAIC EBIC.5 EBIC.25
7852.08719 7983.11043 8051.12157 8017.06459
```

بر طبق برآورد lasso، برآوردهای alasso، enet، scad، mcp را نیز می‌توان به‌دست آورد. حال از برآورد alasso استفاده می‌کنیم (type="lasso").

```
library(regsem)
extractMatrices(out)["A"]
out.reg <- cv_regsem(out, type="lasso", pars_pen=c(1:9), n.lambda=15, jump=.05)
```

یک سلسله دلایل قابل قبول از این‌که چگونه مقادیر زیادی از جریمه را می‌خواهیم از طریق (15 = n.lambda) آزمون کنیم، چگونه مقادیر جریمه برای هر مدل باید افزایش یابد (jump = 0.05)، و در نهایت از برآورد alasso استفاده شده است (type="lasso").

out.reg دو بخش را شامل می‌شود، out.reg[[1]] برآورد پارامتر برای هر یک از 15 مدل است،

```
head(round(out.reg[[1]], 2), 5)
f1 -> x1 f1 -> x2 f1 -> x3 f1 -> x4 f1 -> x5 f1 -> x6 f1 -> x7 f1 -> x8
[1,] 0.51 0.26 0.25 0.98 1.08 0.92 0.2 0.2
[2,] 0.37 0.00 0.00 0.94 1.04 0.86 0.0 0.0
[3,] 0.20 0.00 0.00 0.89 1.00 0.82 0.0 0.0
[4,] 0.00 0.00 0.00 0.86 0.97 0.79 0.0 0.0
[5,] 0.00 0.00 0.00 0.82 0.94 0.76 0.0 0.0
f1 -> x9 1 -> x1 1 -> x2 1 -> x3 1 -> x4 1 -> x5 1 -> x6 1 -> x7 1 -> x8
[1,] 0.31 4.94 6.09 2.25 3.06 4.34 2.19 4.19 5.53
[2,] 0.00 4.94 6.09 2.25 3.06 4.34 2.19 4.19 5.53
[3,] 0.00 4.94 6.09 2.25 3.06 4.34 2.19 4.19 5.53
[4,] 0.00 4.94 6.09 2.25 3.06 4.34 2.19 4.19 5.53
[5,] 0.00 4.94 6.09 2.25 3.06 4.34 2.19 4.19 5.53
1 -> x9 x1 ~~ x1 x2 ~~ x2 x3 ~~ x3 x4 ~~ x4 x5 ~~ x5 x6 ~~ x6 x7 ~~ x7
[1,] 5.37 1.10 1.31 1.21 0.38 0.49 0.36 1.15
[2,] 5.37 1.15 1.37 1.26 0.37 0.44 0.36 1.18
```

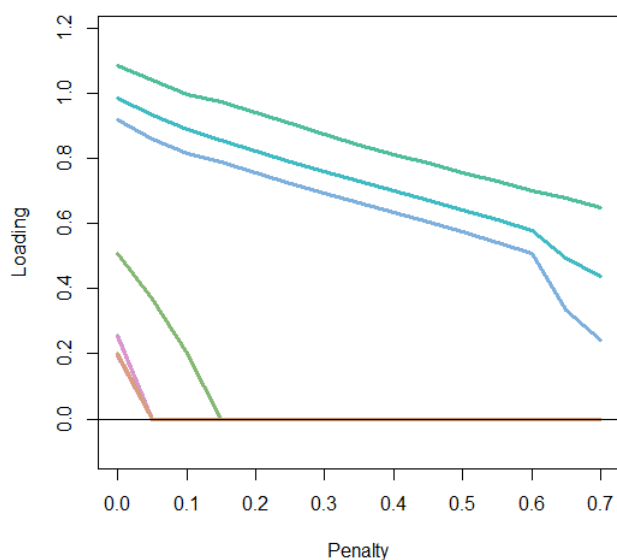
```
[3,] 5.37 1.18 1.38 1.27 0.38 0.44 0.37 1.18
[4,] 5.37 1.28 1.38 1.27 0.39 0.43 0.38 1.18
[5,] 5.37 1.33 1.38 1.27 0.39 0.43 0.38 1.18
x8 ~~ x8 x9 ~~ x9
[1,] 0.98 0.92
[2,] 1.02 1.00
[3,] 1.02 1.01
[4,] 1.02 1.01
[5,] 1.02 1.01
```

در حالی که `out.reg[[2]]` ، اطلاعات مربوط به برازش هر مدل را شامل می شود

```
round(out.reg[[2]],2)
lambda conv rmsea BIC
[1,] 0.00 0 0.19 7856.54
[2,] 0.05 0 0.19 7894.55
[3,] 0.10 0 0.19 7910.66
[4,] 0.15 0 0.20 7941.45
[5,] 0.20 0 0.20 7947.12
[6,] 0.25 0 0.20 7955.17
[7,] 0.30 0 0.21 7964.94
[8,] 0.35 0 0.21 7976.37
[9,] 0.40 0 0.21 7989.57
[10,] 0.45 0 0.22 8004.80
[11,] 0.50 0 0.22 8022.61
[12,] 0.55 0 0.22 8044.11
[13,] 0.60 0 0.23 8070.47
[14,] 0.65 0 0.25 8169.80
[15,] 0.70 0 0.27 8241.56
```

به جای بررسی ماتریس خروجی برآورد پارامترها `out.reg[[1]]` ، کاربران گزینه‌ای برای رسم خط سیر (گذرگاه) هر یک از پارامترهای جریمه دارند.
برای این کار از دستور `plot_cv(out.reg, pars = ۱ : ۹)` استفاده شده است، و در نتیجه داریم

```
plot_cv(out.reg, pars=1:9)
```



شکل ۳.۳: نمودار اثر لاسوی تطبیقی داده‌های بخش ۳.۳

در مدل بالا بهترین جریمه بر اساس شاخص BIC مقدار $\lambda = 0$ می‌باشد. با این حال، برای اثبات این موضوع می‌توانیم مقدار جریمه $0/2$ را انتخاب کنیم. حال مدل را بار دیگر با دستور `regsem()` اجرا می‌کنیم.

```
mod.out <- regsem(out, type="alasso", pars_pen=c(1:9),lambda=0.2)
#summary(mod.out)
```

شاخص‌های برازش اضافی را می‌توان از طریق تابع `fit_indices` به دست آورد.

```
fit_indices(mod.out)
$Data_Type
[1] "Train"

$fits
Fmin varFit p chisq p.chisq
0.72581 0.00000 9.00000 436.93620 0.00000
nfac df npar N baseline.chisq
1.00000 33.00000 21.00000 301.00000 918.85159
baseline.df logl ncp rmsea rmsea.lower
36.00000 -3913.56026 1.34645 0.20199 0.18506
rmsea.upper rmsea.pval CFI TLI BIC
```

```
NA 0.00000 0.54246 0.50087 7946.96984
AIC CAIC EBIC.5 EBIC.25
7869.12053 8015.56462 8087.96358 8051.70937
```

حال از برآورد elastic net استفاده می‌کنیم (type="enet").

```
library(regsem)
extractMatrices(out) ["A"]
out.reg <- cv_regsem(out, type="enet", pars_pen=c(1:9), n.lambda=15, jump=.05)
```

یک سلسله دلایل قابل قبول از این‌که چگونه مقادیر زیادی از جریمه را می‌خواهیم از طریق ($n.lambda = 15$) آزمون کنیم، چگونه مقادیر جریمه برای هر مدل باید افزایش یابد ($jump = 0.05$)، و در نهایت از برآورد *elasticnet* استفاده شده است (type="enet").

out.reg دو بخش را شامل می‌شود، out.reg[[1]] برآورد پارامتر برای هر یک از ۱۵ مدل است،

```
head(round(out.reg[[1]], 2), 5)
f1 -> x1 f1 -> x2 f1 -> x3 f1 -> x4 f1 -> x5 f1 -> x6 f1 -> x7 f1 -> x8
[1,] 0.51 0.26 0.25 0.98 1.08 0.92 0.20 0.20
[2,] 0.49 0.24 0.23 0.96 1.06 0.90 0.18 0.19
[3,] 0.47 0.23 0.22 0.95 1.04 0.88 0.17 0.18
[4,] 0.46 0.22 0.21 0.94 1.03 0.87 0.16 0.17
[5,] 0.45 0.21 0.20 0.92 1.02 0.86 0.15 0.16
f1 -> x9 1 -> x1 1 -> x2 1 -> x3 1 -> x4 1 -> x5 1 -> x6 1 -> x7 1 -> x8
[1,] 0.31 4.94 6.09 2.25 3.06 4.34 2.19 4.19 5.53
[2,] 0.29 4.94 6.09 2.25 3.06 4.34 2.19 4.19 5.53
[3,] 0.28 4.94 6.09 2.25 3.06 4.34 2.19 4.19 5.53
[4,] 0.27 4.94 6.09 2.25 3.06 4.34 2.19 4.19 5.53
[5,] 0.26 4.94 6.09 2.25 3.06 4.34 2.19 4.19 5.53
1 -> x9 x1 ~~ x1 x2 ~~ x2 x3 ~~ x3 x4 ~~ x4 x5 ~~ x5 x6 ~~ x6 x7 ~~ x7
[1,] 5.37 1.10 1.31 1.21 0.38 0.49 0.36 1.15
[2,] 5.37 1.10 1.32 1.21 0.38 0.48 0.36 1.15
[3,] 5.37 1.10 1.32 1.22 0.38 0.48 0.36 1.15
[4,] 5.37 1.11 1.32 1.22 0.38 0.48 0.36 1.15
[5,] 5.37 1.11 1.32 1.22 0.38 0.48 0.36 1.15
x8 ~~ x8 x9 ~~ x9
[1,] 0.98 0.92
```

[2,] 0.98 0.92

[3,] 0.98 0.92

[4,] 0.98 0.93

[5,] 0.98 0.93

در حالی که `out.reg[[2]]` ، اطلاعات مربوط به برازش هر مدل را شامل می‌شود

```
round(out.reg[[2]],2)
```

```
lambda conv rmsea BIC
```

[1,] 0.00 0 0.19 7856.54

[2,] 0.05 0 0.19 7856.99

[3,] 0.10 0 0.19 7857.78

[4,] 0.15 0 0.19 7858.96

[5,] 0.20 0 0.19 7860.31

[6,] 0.25 0 0.19 7862.02

[7,] 0.30 0 0.19 7857.86

[8,] 0.35 0 0.19 7864.06

[9,] 0.40 0 0.19 7858.76

[10,] 0.45 0 0.19 7859.29

[11,] 0.50 0 0.19 7859.87

[12,] 0.55 0 0.19 7860.49

[13,] 0.60 0 0.19 7861.16

[14,] 0.65 0 0.19 7861.88

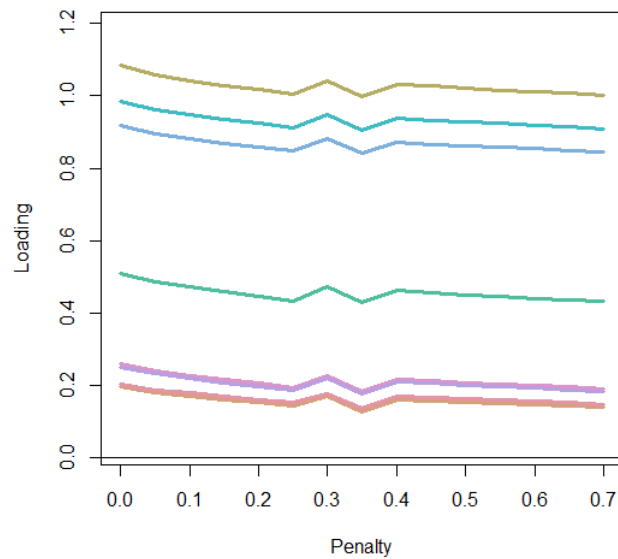
[15,] 0.70 0 0.19 7862.63

به‌جای بررسی ماتریس خروجی برآورد پارامترها `out.reg[[1]]`، کاربران گزینه‌ای برای رسم خط سیر (گذرگاه) هر یک از پارامترهای جریمه دارند.

برای این کار از دستور `plot_cv(out.reg, pars = ۱ : ۹)` استفاده شده است، و در نتیجه

داریم

```
plot_cv(out.reg, pars=1:9)
```



شکل ۴.۳: نمودار اثر الاستیکنت داده‌های بخش ۳.۳

در مدل بالا بهترین جریمه بر اساس شاخص BIC مقدار $\lambda = 0$ می‌باشد. با این حال، برای اثبات این موضوع می‌توانیم مقدار جریمه 0.2 را انتخاب کنیم. حال مدل را بار دیگر با دستور `regsem()` اجرا می‌کنیم.

```
mod.out <- regsem(out, type="enet", pars_pen=c(1:9),lambda=0.2)
#summary(mod.out)
```

شاخص‌های برازش اضافی را می‌توان از طریق تابع `fit_indices` به‌دست آورد.

```
fit_indices(mod.out)
$Data_Type
[1] "Train"

$fits
Fmin varFit p chisq p.chisq
0.72581 0.00000 9.00000 436.93620 0.00000
nfac df npar N baseline.chisq
1.00000 33.00000 21.00000 301.00000 918.85159
baseline.df logl ncp rmsea rmsea.lower
36.00000 -3913.56026 1.34645 0.20199 0.18506
rmsea.upper rmsea.pval CFI TLI BIC
```

```
NA 0.00000 0.54246 0.50087 7946.96984
AIC CAIC EBIC.5 EBIC.25
7869.12053 8015.56462 8087.96358 8051.70937
```

حال از برآورد SCAD استفاده می‌کنیم (type="scad").

```
library(regsem)
extractMatrices(out)["A"]
out.reg <- cv_regsem(out, type="scad", pars_pen=c(1:9), n.lambda=15, jump=.05)
```

یک سلسله دلایل قابل قبول از این‌که چگونه مقادیر زیادی از جریمه را می‌خواهیم از طریق (15 = n.lambda) آزمون کنیم، چگونه مقادیر جریمه برای هر مدل باید افزایش یابد (jump = 0.05)، و در نهایت از برآورد SCAD استفاده شده است (type="scad").

out.reg دو بخش را شامل می‌شود، out.reg[[1]] برآورد پارامتر برای هر یک از 15 مدل است،

```
head(round(out.reg[[1]], 2), 5)
f1 -> x1 f1 -> x2 f1 -> x3 f1 -> x4 f1 -> x5 f1 -> x6 f1 -> x7 f1 -> x8
[1,] 0.51 0.26 0.25 0.98 1.08 0.92 0.20 0.2
[2,] 0.50 0.25 0.24 0.98 1.07 0.91 0.19 0.2
[3,] 0.51 0.26 0.25 0.98 1.08 0.92 0.20 0.2
[4,] 0.51 0.26 0.25 0.98 1.08 0.92 0.20 0.2
[5,] 0.51 0.26 0.25 0.98 1.08 0.92 0.20 0.2
f1 -> x9 1 -> x1 1 -> x2 1 -> x3 1 -> x4 1 -> x5 1 -> x6 1 -> x7 1 -> x8
[1,] 0.31 4.94 6.09 2.25 3.06 4.34 2.19 4.19 5.53
[2,] 0.30 4.94 6.09 2.25 3.06 4.34 2.19 4.19 5.53
[3,] 0.31 4.94 6.09 2.25 3.06 4.34 2.19 4.19 5.53
[4,] 0.31 4.94 6.09 2.25 3.06 4.34 2.19 4.19 5.53
[5,] 0.31 4.94 6.09 2.25 3.06 4.34 2.19 4.19 5.53
1 -> x9 x1 ~~ x1 x2 ~~ x2 x3 ~~ x3 x4 ~~ x4 x5 ~~ x5 x6 ~~ x6 x7 ~~ x7
[1,] 5.37 1.1 1.31 1.21 0.38 0.49 0.36 1.15
[2,] 5.37 1.1 1.32 1.21 0.38 0.48 0.36 1.15
[3,] 5.37 1.1 1.31 1.21 0.38 0.48 0.36 1.15
[4,] 5.37 1.1 1.31 1.21 0.38 0.49 0.36 1.15
[5,] 5.37 1.1 1.31 1.21 0.38 0.48 0.36 1.15
x8 ~~ x8 x9 ~~ x9
[1,] 0.98 0.92
```



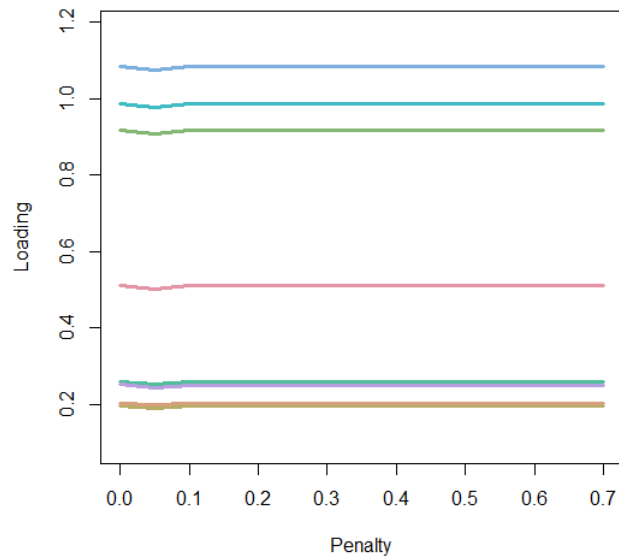
```
[2,] 0.98 0.92
[3,] 0.98 0.92
[4,] 0.98 0.92
[5,] 0.98 0.92
```

در حالی که `out.reg[[2]]` ، اطلاعات مربوط به برازش هر مدل را شامل می‌شود

```
round(out.reg[[2]],2)
lambda conv rmsea BIC
[1,] 0.00 0 0.19 7856.54
[2,] 0.05 0 0.19 7856.61
[3,] 0.10 0 0.19 7856.54
[4,] 0.15 0 0.19 7856.54
[5,] 0.20 0 0.19 7856.54
[6,] 0.25 0 0.19 7856.54
[7,] 0.30 0 0.19 7856.54
[8,] 0.35 0 0.19 7856.54
[9,] 0.40 0 0.19 7856.54
[10,] 0.45 0 0.19 7856.54
[11,] 0.50 0 0.19 7856.54
[12,] 0.55 0 0.19 7856.54
[13,] 0.60 0 0.19 7856.54
[14,] 0.65 0 0.19 7856.54
[15,] 0.70 0 0.19 7856.54
```

به‌جای بررسی ماتریس خروجی برآورد پارامترها `out.reg[[1]]` ، کاربران گزینه‌ای برای رسم خط سیر (گذرگاه) هر یک از پارامترهای جریمه دارند.
برای این کار از دستور `plot_cv(out.reg, pars = ۱ : ۹)` استفاده شده است، و در نتیجه داریم

```
plot_cv(out.reg, pars=1:9)
```



شکل ۵.۳: نمودار اثر scad داده‌های بخش ۳.۳

در مدل بالا بهترین جریمه بر اساس شاخص BIC مقدار $\lambda = 0$ می‌باشد. با این حال، برای اثبات این موضوع می‌توانیم مقدار جریمه $0/2$ را انتخاب کنیم. حال مدل را بار دیگر با دستور `regsem()` اجرا می‌کنیم.

```
mod.out <- regsem(out, type="scad", pars_pen=c(1:9),lambda=0.2)
#summary(mod.out)
```

شاخص‌های برازش اضافی را می‌توان از طریق تابع `fit_indices` به دست آورد.

```
fit_indices(mod.out)
$Data_Type
[1] "Train"

$fits
Fmin varFit p chisq p.chisq
0.51871 0.00000 9.00000 312.26457 0.00000
nfac df npar N baseline.chisq
1.00000 27.00000 27.00000 301.00000 918.85159
baseline.df logl ncp rmsea rmsea.lower
36.00000 -3851.22445 0.95088 0.18766 0.16899
rmsea.upper rmsea.pval CFI TLI BIC
```

```
NA 0.00000 0.67688 0.56918 7856.54088
AIC CAIC EBIC.5 EBIC.25
7756.44890 7856.63043 7915.86594 7886.20341
```

حال از برآورد MCP استفاده می‌کنیم (type="mcp").

```
library(regsem)
extractMatrices(out) ["A"]
out.reg <- cv_regsem(out, type="mcp", pars_pen=c(1:9), n.lambda=15, jump=.05)
```

یک سلسله دلایل قابل قبول از این‌که چگونه مقادیر زیادی از جریمه را می‌خواهیم از طریق ($n.lambda = 15$) آزمون کنیم، چگونه مقادیر جریمه برای هر مدل باید افزایش یابد ($jump = 0.05$)، و در نهایت از برآورد MCP استفاده شده است (type="mcp").

out.reg دو بخش را شامل می‌شود، out.reg[[1]] برآورد پارامتر برای هر یک از ۱۵ مدل است،

```
head(round(out.reg[[1]], 2), 5)
f1 -> x1 f1 -> x2 f1 -> x3 f1 -> x4 f1 -> x5 f1 -> x6 f1 -> x7 f1 -> x8
[1,] 0.51 0.26 0.25 0.98 1.08 0.92 0.2 0.2
[2,] 0.51 0.26 0.25 0.98 1.08 0.92 0.2 0.2
[3,] 0.51 0.26 0.25 0.98 1.08 0.92 0.2 0.2
[4,] 0.51 0.26 0.25 0.98 1.08 0.92 0.2 0.2
[5,] 0.51 0.26 0.25 0.98 1.08 0.92 0.2 0.2
f1 -> x9 1 -> x1 1 -> x2 1 -> x3 1 -> x4 1 -> x5 1 -> x6 1 -> x7 1 -> x8
[1,] 0.31 4.94 6.09 2.25 3.06 4.34 2.19 4.19 5.53
[2,] 0.31 4.94 6.09 2.25 3.06 4.34 2.19 4.19 5.53
[3,] 0.31 4.94 6.09 2.25 3.06 4.34 2.19 4.19 5.53
[4,] 0.31 4.94 6.09 2.25 3.06 4.34 2.19 4.19 5.53
[5,] 0.31 4.94 6.09 2.25 3.06 4.34 2.19 4.19 5.53
1 -> x9 x1 ~~ x1 x2 ~~ x2 x3 ~~ x3 x4 ~~ x4 x5 ~~ x5 x6 ~~ x6 x7 ~~ x7
[1,] 5.37 1.1 1.31 1.21 0.38 0.49 0.36 1.15
[2,] 5.37 1.1 1.31 1.21 0.38 0.49 0.36 1.15
[3,] 5.37 1.1 1.31 1.21 0.38 0.49 0.36 1.15
[4,] 5.37 1.1 1.31 1.21 0.38 0.49 0.36 1.15
[5,] 5.37 1.1 1.31 1.21 0.38 0.49 0.36 1.15
x8 ~~ x8 x9 ~~ x9
[1,] 0.98 0.92
```

```
[2,] 0.98 0.92
[3,] 0.98 0.92
[4,] 0.98 0.92
[5,] 0.98 0.92
```

در حالی که `out.reg[[2]]` ، اطلاعات مربوط به برازش هر مدل را شامل می‌شود

```
round(out.reg[[2]],2)
lambda conv rmsea BIC
[1,] 0.00 0 0.19 7856.54
[2,] 0.05 0 0.19 7856.54
[3,] 0.10 0 0.19 7856.54
[4,] 0.15 0 0.19 7856.54
[5,] 0.20 0 0.19 7856.54
[6,] 0.25 0 0.19 7856.54
[7,] 0.30 0 0.19 7856.60
[8,] 0.35 0 0.19 7856.61
[9,] 0.40 0 0.19 7856.60
[10,] 0.45 0 0.19 7856.61
[11,] 0.50 0 0.19 7856.60
[12,] 0.55 0 0.19 7856.61
[13,] 0.60 0 0.19 7856.60
[14,] 0.65 0 0.19 7856.61
[15,] 0.70 0 0.19 7856.60
```

به‌جای بررسی ماتریس خروجی برآورد پارامترها `out.reg[[1]]`، کاربران گزینه‌ای برای رسم خط سیر (گذرگاه) هر یک از پارامترهای جریمه دارند. برای این کار از دستور `plot_cv(out.reg, pars = ۱ : ۹)` استفاده شده است، و در نتیجه داریم

```
plot_cv(out.reg, pars=1:9)
```

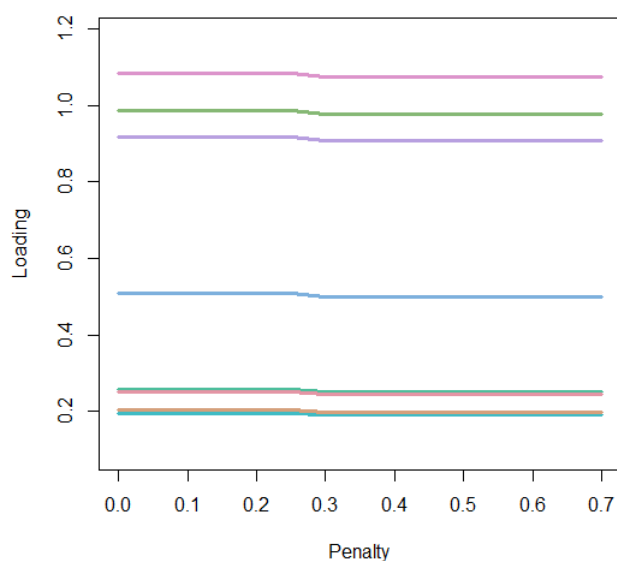
در مدل بالا بهترین جریمه بر اساس شاخص BIC مقدار $\lambda = 0$ می‌باشد. با این حال، برای اثبات این موضوع می‌توانیم مقدار جریمه $0/2$ را انتخاب کنیم. حال مدل را بار دیگر با دستور `regsem()` اجرا می‌کنیم.

```
mod.out <- regsem(out, type="mcp", pars_pen=c(1:9),lambda=0.2)
#summary(mod.out)
```

شاخص‌های برازش اضافی را می‌توان از طریق تابع *fit_indices* به‌دست آورد.

```
fit_indices(mod.out)
$Data_Type
[1] "Train"

$fits
Fmin varFit p chisq p.chisq
0.51871 0.00000 9.00000 312.26425 0.00000
nfac df npar N baseline.chisq
1.00000 27.00000 27.00000 301.00000 918.85159
baseline.df logl ncp rmsea rmsea.lower
36.00000 -3851.22429 0.95088 0.18766 0.16899
rmsea.upper rmsea.pval CFI TLI BIC
NA 0.00000 0.67688 0.56918 7856.54056
AIC CAIC EBIC.5 EBIC.25
7756.44858 7856.63011 7915.86562 7886.20309
```

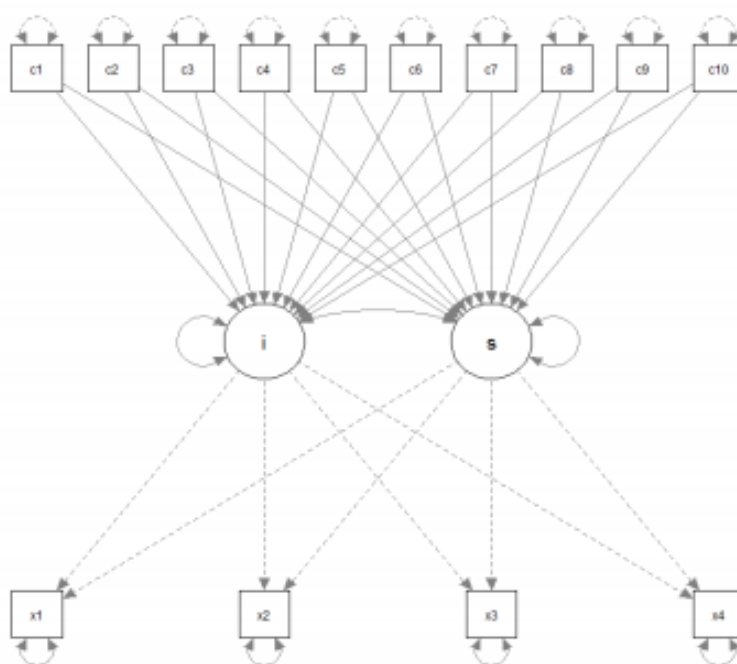


شکل ۶.۳: نمودار اثر mcp داده‌های بخش ۳.۳

مدل سازی معادلات ساختاری بسیار دشوار است و ترکیب شدن (جفت شدن) با تنظیم باعث ساده تر شدن آن نمی شود.

۴.۳ مقایسه

برای مقایسه انواع جریمه در regsem، به مثال مدل منحنی رشد پنهان نمایش داده شده در شکل زیر می پردازیم.



شکل ۷.۳: نمودار مدل منحنی رشد پنهان

با استفاده از داده های شبیه سازی شده، مدل را می توان در بسته lavaan اجرا کرد.

```
mod1 <- "
i =~ 1*x1 + 1*x2 + 1*x3 + 1*x4
s =~ 0*x1 + 1*x2 + 2*x3 + 3*x4
i ~ c1 + c2 + c3 + c4 + c5 + c6 + c7 + c8 + c9 + c10
s ~ c1 + c2 + c3 + c4 + c5 + c6 + c7 + c8 + c9 + c10
"
lav.growth <- growth(mod1,dat,fixed.x=T)
```

مقایسه انواع مختلف جریمه‌ها در بسته regsem نیاز به مشخصات دیگری از انواع آرگومان (توابع) دارد. گزینه‌های موجود شامل روش‌های ماکزیمم درست‌نمایی ("none")، ریج ("ridge")، لاسو ("lasso")، لاسوی تطبیقی ("alasso")، الاستیک‌نت ("enet")، جریمه مقعر اسکد ("scad") و جریمه مقعر مینیماکس ("mcp") می‌باشد. برای جریمه الاستیک‌نت، یک پارامتر اضافی وجود دارد، α بین روش ریج و لاسو تعادل برقرار می‌کند که به‌طور پیش‌فرض $\alpha = 0.5$ را اختیار می‌کند. علاوه بر روش الاستیک‌نت، روش‌های جریمه مقعر اسکد و جریمه مقعر مینیماکس نیز پارامتر اضافی دارند، γ که به‌طور پیش‌فرض $\gamma = 3/7$ را اختیار می‌کند (فن و لی^{۱۹}، ۲۰۰۱).

برای مقایسه کردن هر یک از ۲۰ متغیر کمکی رگرسیون را با استفاده از جریمه‌های ریج، لاسو، لاسوی تطبیقی، الاستیک‌نت، جریمه مقعر اسکد و جریمه مقعر مینیماکس با برآورد ماکزیمم درست‌نمایی مقایسه می‌کنیم.

در این مدل، داده‌ها برای دو اثر بزرگ (هر دو پارامتر c_1)، دو اثر کوچک (هر دو پارامتر c_2) و ۱۶ تا اثر صفر (پارامترهای c_3 تا c_{10}) شبیه‌سازی شده‌اند. توجه داشته باشید که هدف از شبیه‌سازی متغیرهای کمکی، وجود کواریانس صفر در میان هر متغیر می‌باشد. اگر همبستگی قابل ملاحظه‌ای در میان متغیرهای کمکی بود، جریمه الاستیک‌نت به‌طور همزمان در انتخاب عوامل پیش‌بینی‌کننده و هم‌چنین محاسبه همگرایی به‌طور همزمان مناسب‌تر خواهد بود. برآورد پارامتر متناظر با بهترین برازش BIC، مناسب هر مدل هستند. نتایج جدول ۷.۲ با استفاده از بسته xtable به‌دست آمده است (داهل^{۲۰}، ۲۰۰۹).

^{۱۹}Fan and Li

^{۲۰}Dahl

جدول ۱.۳: برآورد پارامتر برای مدل‌های نهایی در ۵ روش برآورد

	<i>MLE</i>	<i>Lasso</i>	<i>ALasso</i>	<i>Scad</i>	<i>Mcp</i>
$c1- > i$	۰/۹۲*	۰/۷۲	۰/۹۱	۰/۹۴	۰/۹۲
$c2- > i$	۰/۰۷	۰/۰۰	۰/۰۰	۰/۰۰	۰/۰۰
$c3- > i$	۰/۱۰	۰/۰۰	۰/۰۰	۰/۰۰	۰/۰۰
$c4- > i$	۰/۰۷	۰/۰۰	۰/۰۰	۰/۰۰	۰/۰۰
$c5- > i$	۰/۰۴	۰/۰۰	۰/۰۰	۰/۰۰	۰/۰۰
$c6- > i$	-۰/۲۵	۰/۰۰	۰/۰۰	۰/۰۰	-۰/۱۹
$c7- > i$	۰/۱۱	۰/۰۰	۰/۰۰	۰/۰۰	۰/۰۰
$c8- > i$	-۰/۱۳	۰/۰۰	۰/۰۰	۰/۰۰	۰/۰۰
$c9- > i$	-۰/۰۳	۰/۰۰	۰/۰۰	۰/۰۰	۰/۰۰
$c10- > i$	۰/۰۹	۰/۰۰	۰/۰۰	۰/۰۰	۰/۰۰
$c1- > s$	۱/۱۸*	۱/۰۹	۱/۲۲	۱/۲۴	۱/۲۴
$c2- > s$	۰/۲۹*	۰/۱۹	۰/۲۸	۰/۳۵	۰/۳۵
$c3- > s$	۰/۱۸	۰/۰۹	۰/۰۰	۰/۰۰	۰/۰۰
$c4- > s$	-۰/۰۸	۰/۰۰	۰/۰۰	۰/۰۰	۰/۰۰
$c5- > s$	-۰/۱۸	۰/۰۰	۰/۰۰	۰/۰۰	۰/۰۰
$c6- > s$	۰/۲۵*	۰/۰۰	۰/۰۰	۰/۰۰	-۰/۱۹
$c7- > s$	-۰/۱۸	-۰/۰۴	۰/۰۰	۰/۰۰	۰/۰۰
$c8- > s$	۰/۲۶*	۰/۰۰	۰/۰۰	۰/۰۰	۰/۰۰
$c9- > s$	-۰/۰۶	۰/۰۰	۰/۰۰	۰/۰۰	۰/۰۰
$c10- > s$	۰/۰۸	۰/۰۰	۰/۰۰	۰/۰۰	۰/۰۰
<i>BIC</i>	۳۴۶۵/۲۸	۳۴۲۷/۴۶	۳۴۱۵/۰۵	۳۴۱۴/۳۸	۳۴۱۷/۲۰

توجه داشته باشید که * نشان دهنده سطح معناداری پارامترهای با اهمیت در سطح ۰/۰۵ برای برآورد ماکزیمم درست‌نمایی (ML) می‌باشد.

۵.۳ نتیجه‌گیری

در این فصل یک مرور کلی روی استفاده از بسته regsem به عنوان مدل‌سازی معادلات ساختاری تنظیم ارائه کردیم. از آنجا که مدل‌سازی معادلات ساختاری یک طیف گسترده‌ای از مدل متغیرهای پنهان را شامل می‌شود، بسته regsem، به عنوان یک بسته کلی (عمومی) برای ایجاد فرم‌های (شکل‌های) مختلف تنظیم نسبت به یک گروه از مدل‌ها با متغیرهای پنهان ایجاد شد. بنابراین Regsem و بسته regsem، بر روی طیف گسترده‌ای از مدل‌های SEM شامل تحلیل عاملی تائیدی (جاکوبوچی و همکاران^{۲۱}، ۲۰۱۶)، مدل‌های تغییر نمرات (امتیازات) پنهان (جاکوبوچی و گریم^{۲۲}، ۲۰۱۷) و مدل‌های میانجی (سرنگ^{۲۳} و همکاران در مطبوعات ارزیابی شده است. در آینده بسته regsem به‌روزرسانی می‌شود تا بر روی کاهش زمان محاسبات مدل‌های متغیرهای پنهان بزرگ به‌منظور فراهم آوردن آزمون‌هایی برای جمع‌آوری مجموعه داده‌های بزرگ و بزرگ‌تر برای محققان تمرکز خواهد کرد.

Regsem روشی است که در تمام سطوح طیف داده‌ها عمل می‌کند. روش Regsem زمانی که حجم نمونه کوچک است، علاوه بر کاهش بعد در حضور داده‌های بزرگ (هر دوی N و P) باعث می‌شود پیچیدگی مدل کاهش یابد.

^{۲۱}Jacobucci et al

^{۲۲}Jacobucci and Grimm

^{۲۳}Serang

خلاصه و پیشنهادات برای آینده تحقیق

۶.۳ خلاصه

در این پایان‌نامه مدل معادلات ساختاری را معرفی کرده و جوانب مختلف آن را چه از منظر نظری و چه از منظر کاربردی برای مدل‌بندی متغیرهایی که نمی‌توان آن‌ها را به‌طور مستقیم اندازه‌گیری کرد، مورد بحث و بررسی قرار دادیم. در این راستا مدل‌های اندازه‌گیری و ساختاری را به‌طور جداگانه و توأم ارائه کرده در قالب یک مسئله با وجود داده‌های هم‌خط مورد مطالعه قرار دادیم. از آنجایی که با وجود داده‌های هم‌خط روش‌های کلاسیک از قبیل کمترین توان‌های دوم عملکرد مطلوبی ندارند، برای تحلیل متغیرهای پنهان، مدل معادلات ساختاری ریج را با دو تکنیک متفاوت تحلیل کردیم. در یک تکنیک، ساختار برآوردگر رگرسیون ریج را به‌طور مستقیم بر روی ماتریس همبستگی و کواریانس پیاده‌سازی کردیم. در تکنیکی دیگر ایده روش‌های جریمه شده را برای مدل معادلات ساختاری مورد استفاده قرار دادیم. کاربرد و چگونگی رویارویی با این روش‌ها، در قالب مثال‌های شبیه‌سازی شده و واقعی مورد بررسی قرار گرفت.

۷.۳ پیشنهادات برای آینده تحقیق

با توجه به تحقیقات انجام شده و نتایجی که از تحلیل روش‌ها به‌دست آوردیم می‌توان پیشنهادات زیر را برای تحقیقات آتی ارائه کرد.

۱- از آنجایی که ساختار برآوردگرهای ریج نسبت به پارامتر آن غیرخطی است می‌توان از روش جایگزین لیو به جای ریج استفاده کرد. در این راستا پیشنهاد می‌شود برآوردگر ماتریس کواریانس محاسبه شده و سپس برآوردگر نوع لیو آن محاسبه شود.

۲- در این پایان‌نامه دیدگاه بیزی معادلات ساختاری مورد بررسی قرار نگرفت. می‌توان روش ریج را برای معادلات ساختاری بیزی به کار برد.

۳- برخی از متغیرهایی که در مطالعات عددی به کار بردیم دارای ماهیت غیرخطی در رابطه با متغیر پاسخ هستند. به منظور افزایش کارایی معادلات ساختاری ریج، پیشنهاد می شود مدل نیمه پارامتری به کار رود. در این راستا، متغیرهایی که رابطه غیرخطی با متغیر پاسخ در مدل اندازه گیری دارند را به صورت ناپارامتری وارد مدل می کنیم.

۴- همان طور که در فصل ۳ مشاهده شد برخی روش های جریمه شده در مدل معادلات ساختاری، علاوه بر ریج، مورد بررسی قرار گرفتند. در ادامه این کار می توان از روش های دیگری از قبیل لاسوی گروهی، الاستیک نت گروهی و SCAD گروهی برای متغیرهایی که ساختار گروهی دارند استفاده کرد.

۵- در پایان، می توان بر روی روش های برآورد پارامتر ریج در مدل معادلات ساختاری تحقیق کرد.

پیوست آ

خروجی نرم افزارها

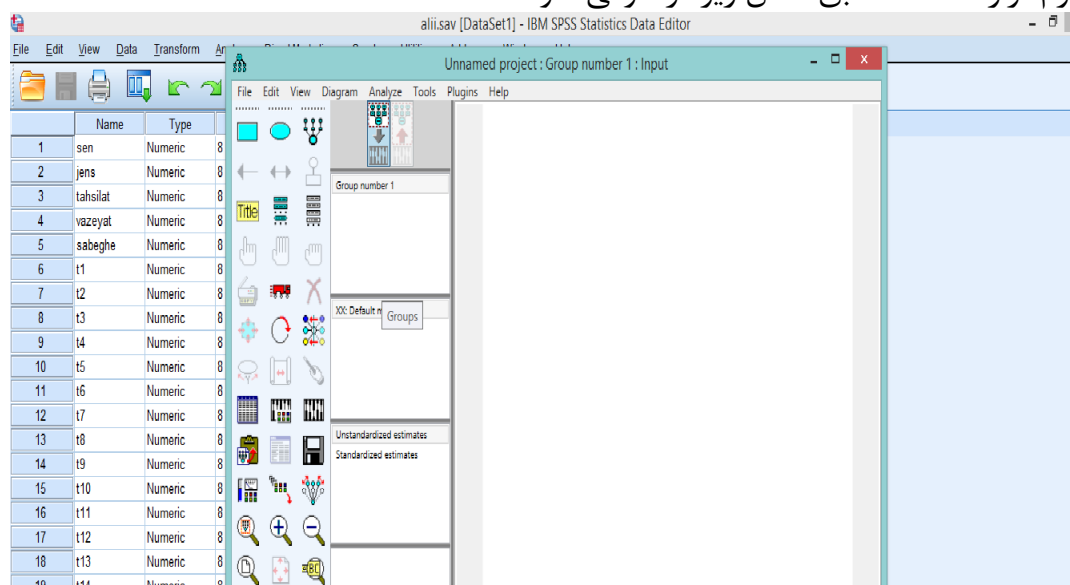
۱.آ روش اجرای مدل سازی معادلات ساختاری در نرم افزار Amos برای مثال مدیریت ورزشی در فصل ۱

ابتدا داده ها را در نرم افزار Spss مطابق شکل زیر وارد کرده

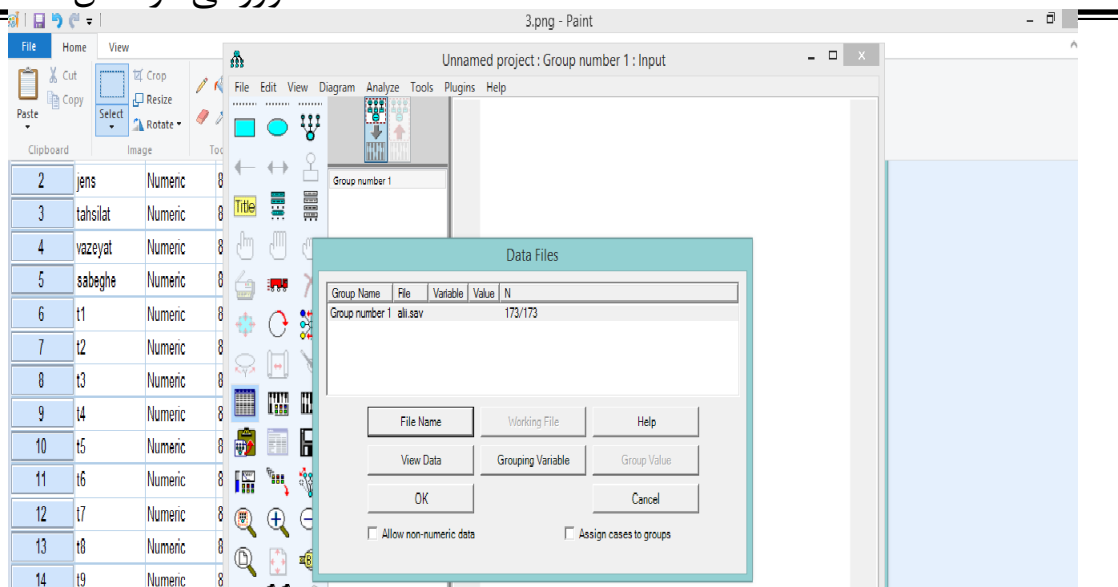
	sen	jens	tahsilat	vazeh	sabeghe	t1	t2	t3	t4	t5	t6	t7	t8	t9	t10	t11	t12	t13	t14	t15	t16	t17	t18	r19	r20	r21	s22
1	3	1	3	3	1	5	5	5	5	5	5	5	5	5	5	5	1	3	5	5	5	5	5	5	3	3	4
2	3	1	4	1	2	5	5	5	5	5	4	5	5	5	5	4	2	2	4	3	4	3	2	2	2	4	2
3	3	1	2	1	2	4	4	5	4	4	4	5	4	5	5	5	2	2	2	3	4	4	4	4	4	4	3
4	4	2	1	1	4	3	4	5	1	5	4	5	3	5	5	4	3	2	4	2	4	3	1	1	1	1	2
5	3	1	4	3	3	2	3	2	3	3	2	3	3	3	3	4	4	3	4	2	1	3	3	3	2	3	1
6	3	2	4	2	2	3	3	2	3	3	3	2	3	3	2	3	3	3	3	2	3	3	3	2	3	3	3
7	3	1	1	1	2	3	3	3	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
8	3	1	4	1	2	3	2	3	3	3	3	3	3	4	3	3	4	3	3	2	4	3	4	3	3	3	2
9	4	2	4	3	5	3	4	3	1	1	2	3	2	4	3	4	4	2	3	2	2	3	2	2	2	2	4
10	3	1	4	2	1	3	4	2	2	4	2	3	4	2	2	4	2	3	2	2	3	2	2	2	2	3	2
11	2	2	2	2	2	5	5	5	5	5	4	4	4	4	4	4	4	5	5	5	4	4	4	4	4	5	4
12	3	1	3	3	2	5	5	5	5	4	5	5	5	5	5	5	4	1	4	4	5	5	4	3	4	4	3
13	4	1	3	3	3	4	4	3	2	2	4	4	4	4	4	4	3	2	2	4	4	4	4	2	2	2	4
14	3	2	3	3	2	4	4	4	4	3	4	4	4	4	4	4	3	1	4	3	4	4	2	3	4	5	2
15	3	2	4	1	2	3	4	3	2	1	4	4	4	4	4	3	1	2	4	4	4	4	4	4	2	2	4
16	3	2	4	1	1	3	4	4	3	4	2	1	4	4	3	4	1	2	2	1	4	2	2	2	2	2	3

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	sen	Numeric	8	0		{1, KAMTA...	None	4	Center	Ordinal	Input
2	jens	Numeric	8	0		{1, MARJ...	None	4	Center	Nominal	Input
3	tahsilat	Numeric	8	0		{1, DIPLO...	None	4	Center	Nominal	Input
4	vazeyat	Numeric	8	0		{1, GHARA...	None	4	Center	Nominal	Input
5	sabeghe	Numeric	8	0		{1, ZIRE 5 S...	None	4	Center	Ordinal	Input
6	t1	Numeric	8	0		{1, KAMLA...	None	4	Center	Scale	Input
7	t2	Numeric	8	0		{1, KAMLA...	None	4	Center	Scale	Input
8	t3	Numeric	8	0		{1, KAMLA...	None	4	Center	Scale	Input
9	t4	Numeric	8	0		{1, KAMLA...	None	4	Center	Scale	Input
10	t5	Numeric	8	0		{1, KAMLA...	None	4	Center	Scale	Input
11	t6	Numeric	8	0		{1, KAMLA...	None	4	Center	Scale	Input
12	t7	Numeric	8	0		{1, KAMLA...	None	4	Center	Scale	Input
13	t8	Numeric	8	0		{1, KAMLA...	None	4	Center	Scale	Input
14	t9	Numeric	8	0		{1, KAMLA...	None	4	Center	Scale	Input
15	t10	Numeric	8	0		{1, KAMLA...	None	4	Center	Scale	Input
16	t11	Numeric	8	0		{1, KAMLA...	None	4	Center	Scale	Input
17	t12	Numeric	8	0		{1, KAMLA...	None	4	Center	Scale	Input
18	t13	Numeric	8	0		{1, KAMLA...	None	4	Center	Scale	Input

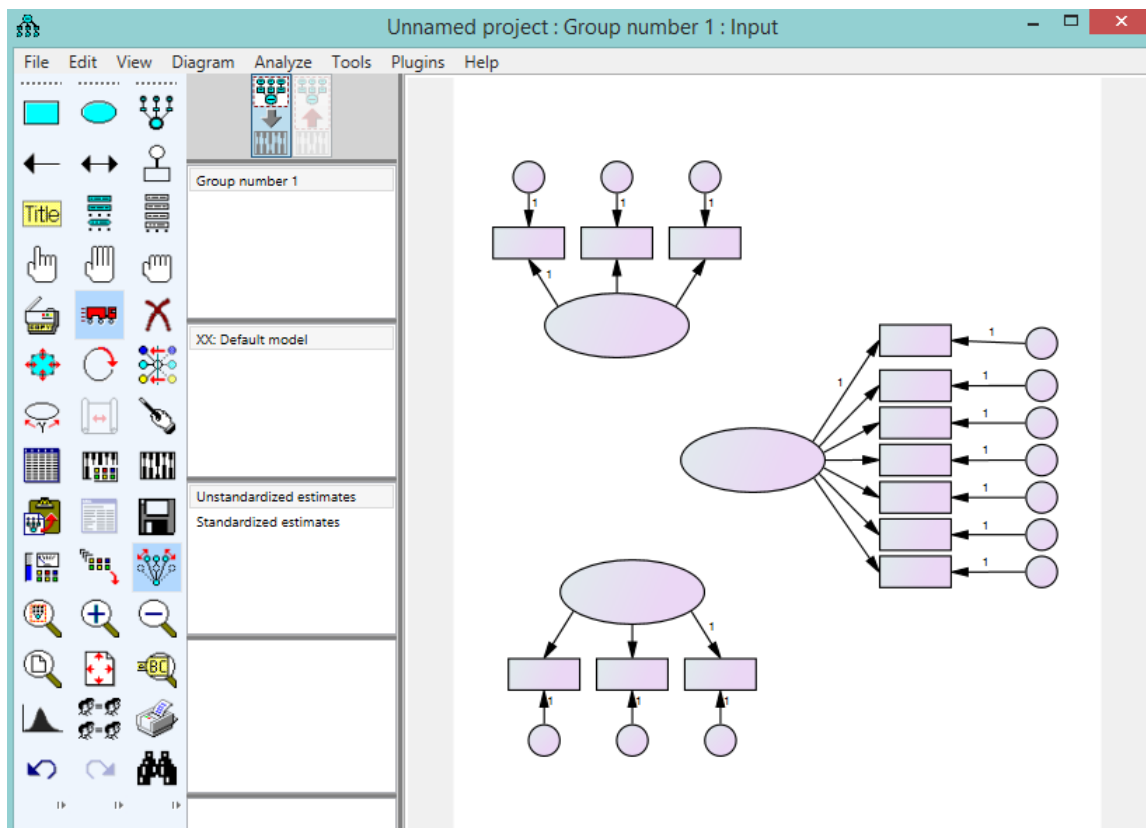
سپس نرم افزار Amos را باز کرده و از منوی File گزینه File Name را انتخاب و داده های موجود در فایل Spss را فراخوانی می کنیم و سپس بر روی گزینه ی Ok کلیک می کنیم تا داده ها در نرم افزار Amos مطابق شکل زیر فراخوانی شوند.



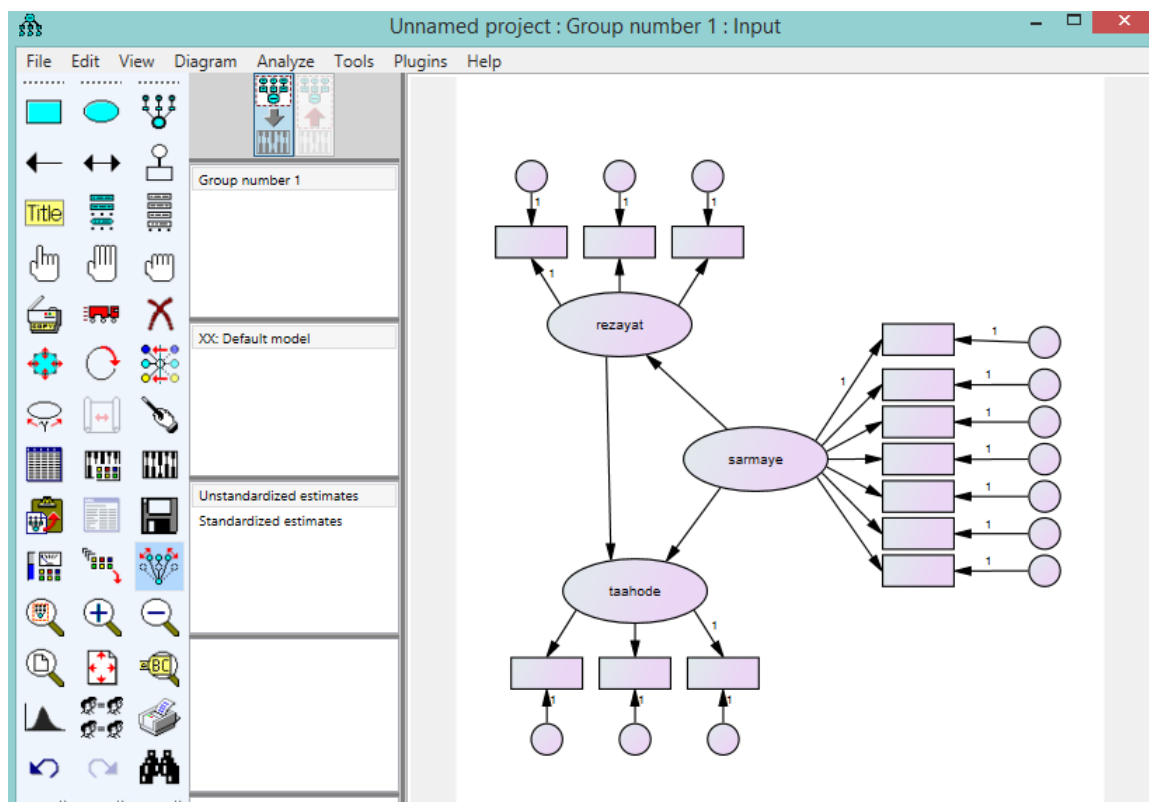
روش اجرای مدل سازی معادلات ساختاری در نرم افزار Amos برای مثال مدیریت ورزشی در فصل ۱ ۱۱۳



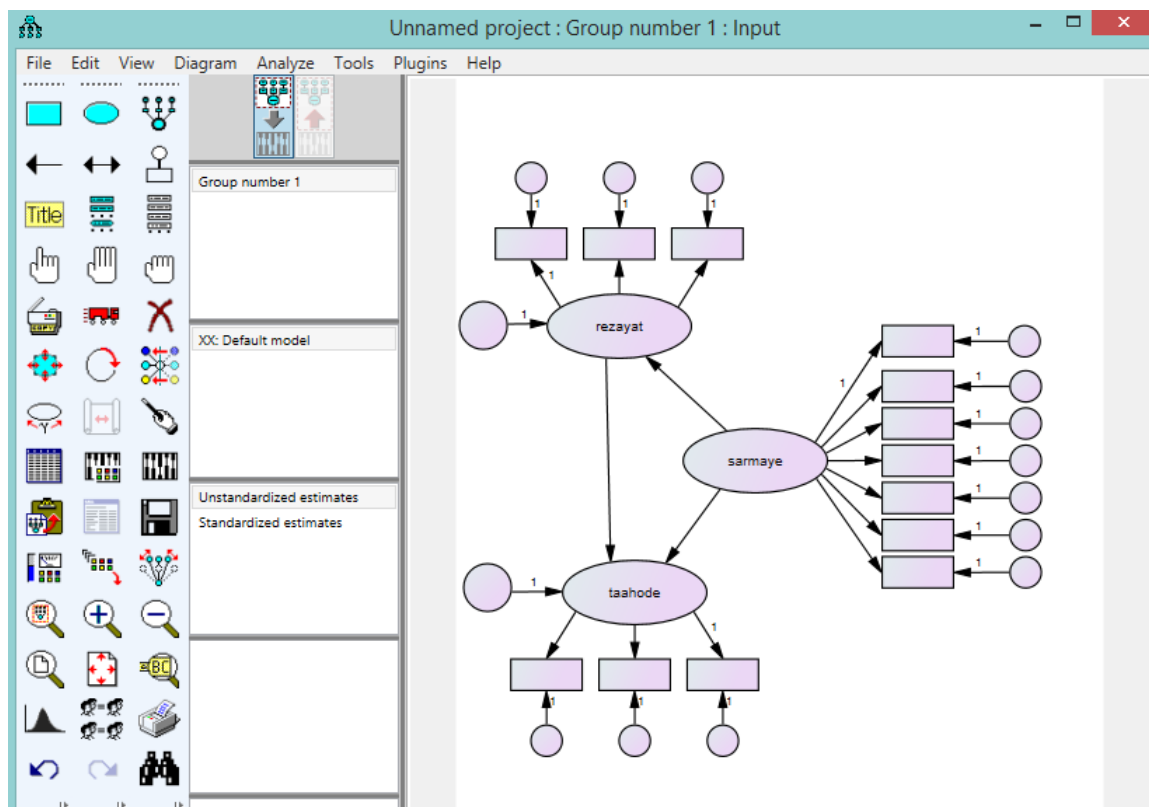
حال با استفاده از ابزار نرم افزار Amos شروع به رسم گراف (شکل) می کنیم.



در این جا ضرایب رگرسیونی بین متغیرهای پنهان را نیز رسم می کنیم.

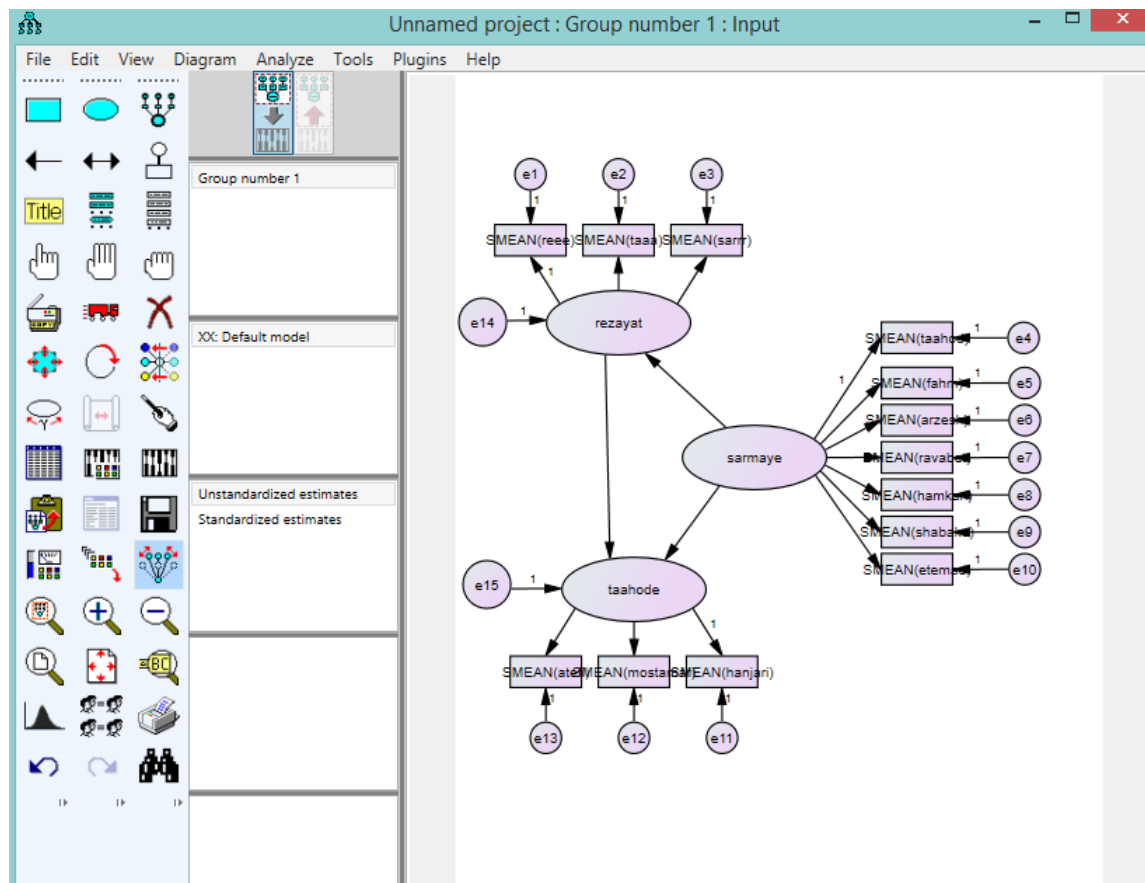


سپس خطاهای متغیرهای پنهان درون را نیز مطابق شکل زیر رسم می کنیم

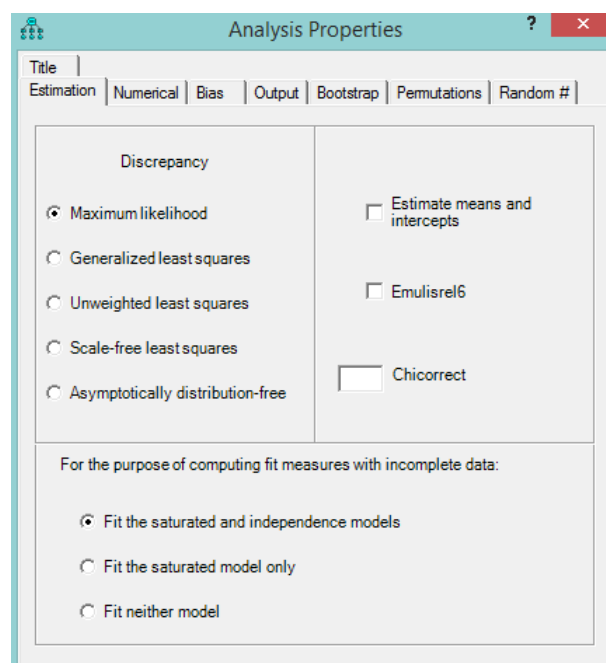


روش اجرای مدل سازی معادلات ساختاری در نرم افزار Amos برای مثال مدیریت ورزشی در فصل ۱ ۱۱۵

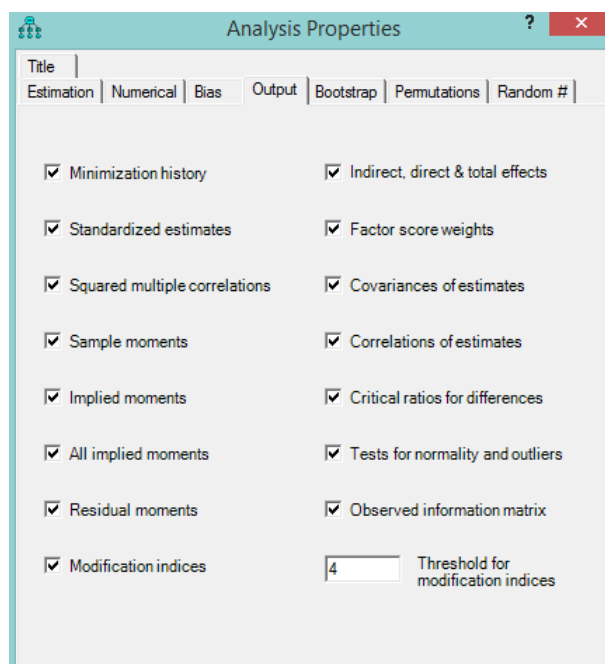
و در آخر نام متغیرهای آشکار (مشاهده شده) را درون مستطیل ها و نام خطاها را درون دایره ها (بیضی) می نویسیم.



حال از منوی View روی گزینه Analysis Properties کلیک کرده و منوی زیر ظاهر می شود

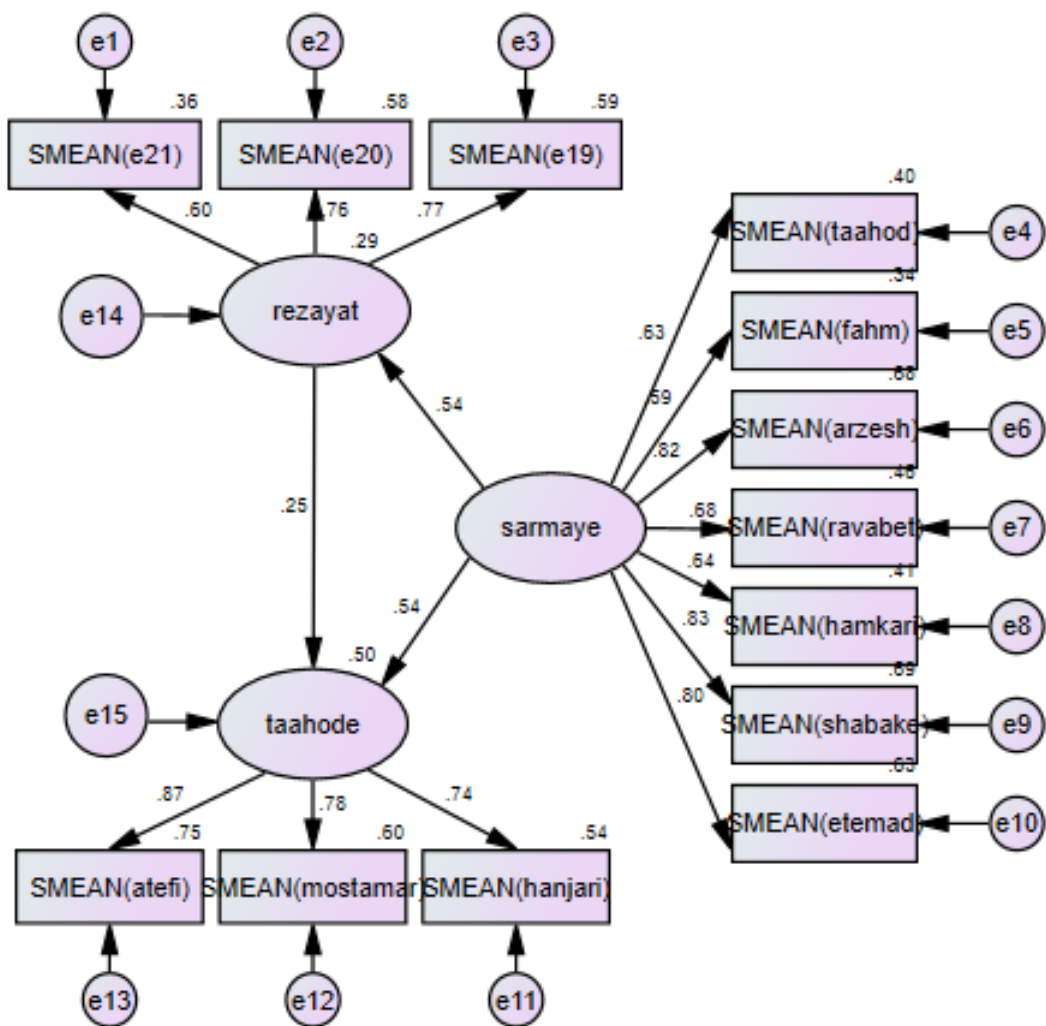


در این جا باید از منوی Estimation گزینه‌ی Estimate means and intercept را علامت زده و حال روی گزینه‌ی Out put کلیک کرده و مواردی را که نیاز داریم علامت بزنییم



پس از ذخیره کردن فایل حال از طریق منوی Analyze روی گزینه‌ی Calculate Estimates کلیک کرده تا برنامه اجرا شود

روش اجرای مدل سازی معادلات ساختاری در نرم افزار Amos برای مثال مدیریت ورزشی در فصل ۱ ۱۱۷



Model Fit Summary

CMIN

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	42	121.751	62	.000	1.964
Saturated model	104	.000	0		
Independence model	26	1107.480	78	.000	14.198

Baseline Comparisons

Model	NFI	RFI	IFI	TLI	CFI
	Delta1	rho1	Delta2	rho2	
Default model	.890	.862	.943	.927	.942
Saturated model	1.000		1.000		1.000
Independence model	.000	.000	.000	.000	.000

Parsimony-Adjusted Measures

Model	PRATIO	PNFI	PCFI
Default model	.795	.707	.749
Saturated model	.000	.000	.000
Independence model	1.000	.000	.000

روش اجرای مدل سازی معادلات ساختاری در نرم افزار Amos برای مثال مدیریت ورزشی در فصل ۱ ۱۱۹
 = NCP

Model	NCP	LO 90	HI 90
Default model	59.751	32.239	95.056
Saturated model	.000	.000	.000
Independence model	1029.480	925.618	1140.761

FMIN

Model	FMIN	F0	LO 90	HI 90
Default model	.708	.347	.187	.553
Saturated model	.000	.000	.000	.000
Independence model	6.439	5.985	5.382	6.632

RMSEA

Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	.075	.055	.094	.022
Independence model	.277	.263	.292	.000

AIC

Model	AIC	BCC	BIC	CAIC
Default model	205.751	213.194		
Saturated model	208.000	226.430		
Independence model	1159.480	1164.088		

ECVI

Model	ECVI	LO 90	HI 90	MECVI
Default model	1.196	1.036	1.401	1.240
Saturated model	1.209	1.209	1.209	1.316
Independence model	6.741	6.137	7.388	6.768

HOELTER

Model	HOELTER .05	HOELTER .01
Default model	115	129
Independence model	16	18

روش اجرای مدل سازی معادلات ساختاری در نرم افزار Amos برای مثال مدیریت ورزشی در فصل ۱ ۱۲۱

Estimates (Group number 1 - Default model)

Scalar Estimates (Group number 1 - Default model)

Maximum Likelihood Estimates

Regression Weights: (Group number 1 - Default model)

			Estimate	S.E.	C.R.	P	Label
rezayat	<---	sarmaye	.854	.188	4.533	***	par_12
taahode	<---	rezayat	.214	.083	2.577	.010	par_5
taahode	<---	sarmaye	.722	.148	4.880	***	par_13
fahm_1	<---	sarmaye	1.067	.159	6.732	***	par_1
arzesh_1	<---	sarmaye	1.679	.192	8.750	***	par_2
e20_1	<---	rezayat	1.324	.201	6.581	***	par_3
mostamar_1	<---	taahode	.943	.098	9.648	***	par_4
hanjari_1	<---	taahode	1.000				
atefi_1	<---	taahode	1.308	.132	9.884	***	par_6
e19_1	<---	rezayat	1.353	.215	6.299	***	par_7
e21_1	<---	rezayat	1.000				
shabake_1	<---	sarmaye	1.752	.203	8.635	***	par_8
hamkari_1	<---	sarmaye	1.880	.260	7.235	***	par_9
ravabet_1	<---	sarmaye	1.016	.133	7.656	***	par_10
etemad_1	<---	sarmaye	1.526	.182	8.401	***	par_11
taahod_1	<---	sarmaye	1.000				

Standardized Regression Weights: (Group number 1 - Default model)

			Estimate
rezayat	<---	sarmaye	.540
taahode	<---	rezayat	.253
taahode	<---	sarmaye	.540
fahm_1	<---	sarmaye	.585
arzesh_1	<---	sarmaye	.822
e20_1	<---	rezayat	.761
mostamar_1	<---	taahode	.777
hanjari_1	<---	taahode	.737
atefi_1	<---	taahode	.866
e19_1	<---	rezayat	.768
e21_1	<---	rezayat	.599
shabake_1	<---	sarmaye	.828
hamkari_1	<---	sarmaye	.644
ravabet_1	<---	sarmaye	.680
etemad_1	<---	sarmaye	.796
taahod_1	<---	sarmaye	.632

Covariances: (Group number 1 - Default model)

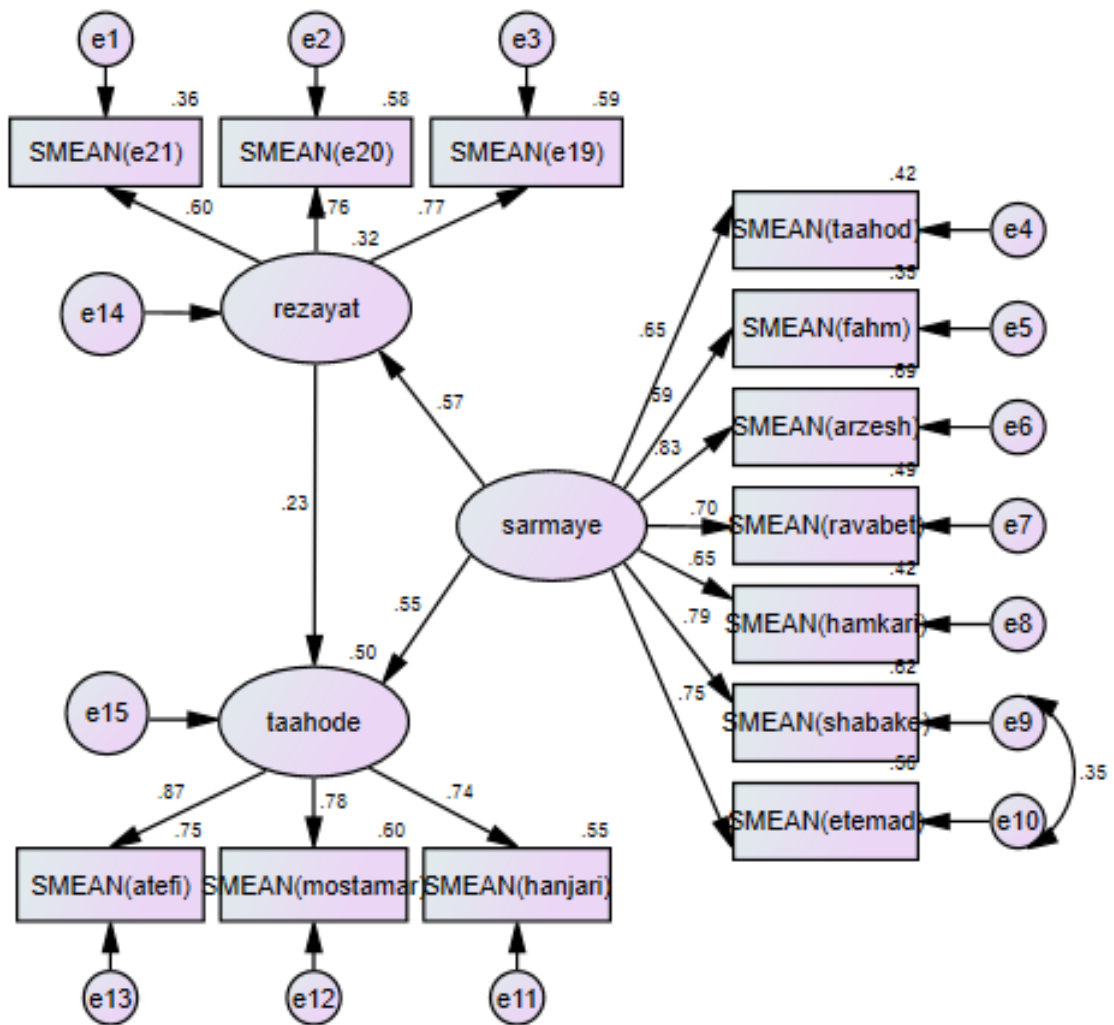
	M.I.	Par Change
e2 <--> e15	4.082	-.050
e2 <--> e13	4.488	-.061
e1 <--> e15	8.409	.077
e1 <--> e13	26.028	.156
e1 <--> e3	4.261	-.100
e9 <--> e14	5.164	-.051
e9 <--> e10	11.647	.059
e7 <--> e9	4.020	-.032
e6 <--> e13	4.549	-.038
e5 <--> e3	5.780	-.081
e5 <--> e2	4.567	.071
e4 <--> e7	4.385	.032

Variances: (Group number 1 - Default model)

	M.I.	Par Change
--	------	------------

Regression Weights: (Group number 1 - Default model)

	M.I.	Par Change
e21_1 <--- taahode	7.081	.366



روش اجرای مدل سازی معادلات ساختاری در نرم افزار Amos برای مثال مدیریت ورزشی در فصل ۱ ۱۲۵

Model Fit Summary

CMIN

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	43	106.960	61	.000	1.753
Saturated model	104	.000	0		
Independence model	26	1107.480	78	.000	14.198

Baseline Comparisons

Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Default model	.903	.877	.956	.943	.955
Saturated model	1.000		1.000		1.000
Independence model	.000	.000	.000	.000	.000

Parsimony-Adjusted Measures

Model	PRATIO	PNFI	PCFI
Default model	.782	.707	.747
Saturated model	.000	.000	.000
Independence model	1.000	.000	.000

NCP

Model	NCP	LO 90	HI 90
Default model	45.960	21.048	78.726
Saturated model	.000	.000	.000
Independence model	1029.480	925.618	1140.761

FMIN

Model	FMIN	F0	LO 90	HI 90
Default model	.622	.267	.122	.458
Saturated model	.000	.000	.000	.000
Independence model	6.439	5.985	5.382	6.632

RMSEA

Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	.066	.045	.087	.100
Independence model	.277	.263	.292	.000

روش اجرای مدل سازی معادلات ساختاری در نرم افزار Amos برای مثال مدیریت ورزشی در فصل ۱ ۱۲۷

AIC

Model	AIC	BCC	BIC	CAIC
Default model	192.960	200.580		
Saturated model	208.000	226.430		
Independence model	<u>1159.480</u>	1164.088		

ECVI

Model	ECVI	LO 90	HI 90	MECVI
Default model	1.122	.977	1.312	1.166
Saturated model	1.209	1.209	1.209	1.316
Independence model	6.741	6.137	7.388	6.768

HOELTER

Model	HOELTER .05	HOELTER .01
Default model	130	145
Independence model	16	18

Estimates (Group number 1 - Default model)

Scalar Estimates (Group number 1 - Default model)

Maximum Likelihood Estimates

Regression Weights: (Group number 1 - Default model)

			Estimate	S.E.	C.R.	P	Label
rezayat	<---	sarmaye	.871	.187	4.669	***	par_12
taahode	<---	rezayat	.198	.085	2.319	.020	par_5
taahode	<---	sarmaye	.716	.148	4.825	***	par_13
fahm_1	<---	sarmaye	1.053	.153	6.886	***	par_1
arzesh_1	<---	sarmaye	1.650	.184	8.951	***	par_2
e20_1	<---	rezayat	1.319	.200	6.593	***	par_3
mostamar_1	<---	taahode	.941	.098	9.650	***	par_4
hanjari_1	<---	taahode	1.000				
atefi_1	<---	taahode	1.304	.132	9.879	***	par_6
e19_1	<---	rezayat	1.348	.213	6.322	***	par_7
e21_1	<---	rezayat	1.000				
shabake_1	<---	sarmaye	1.615	.190	8.485	***	par_8
hamkari_1	<---	sarmaye	1.833	.250	7.337	***	par_9
ravabet_1	<---	sarmaye	1.015	.128	7.955	***	par_10
etemad_1	<---	sarmaye	1.393	.171	8.169	***	par_11
taahod_1	<---	sarmaye	1.000				

Standardized Regression Weights: (Group number 1 - Default model)

		Estimate
rezayat	<--- sarmaye	.566
taahode	<--- rezayat	.234
taahode	<--- sarmaye	.551
fahm_1	<--- sarmaye	.595
arzesh_1	<--- sarmaye	.832
e20_1	<--- rezayat	.760
mostamar_1	<--- taahode	.777
hanjari_1	<--- taahode	.739
atefi_1	<--- taahode	.865
e19_1	<--- rezayat	.767
e21_1	<--- rezayat	.601
shabake_1	<--- sarmaye	.785
hamkari_1	<--- sarmaye	.646
ravabet_1	<--- sarmaye	.700
etemad_1	<--- sarmaye	.748
taahod_1	<--- sarmaye	.651

Modification Indices (Group number 1 - Default model)**Covariances: (Group number 1 - Default model)**

	M.I.	Par Change
e2 <--> e15	4.012	-.050
e2 <--> e13	4.460	-.060
e1 <--> e15	7.957	.075
e1 <--> e13	25.675	.155
e1 <--> e3	4.374	-.102
e6 <--> e15	4.315	-.032
e6 <--> e13	5.250	-.041
e6 <--> e9	4.714	.037
e5 <--> e3	6.467	-.085
e5 <--> e2	4.131	.068

Variances: (Group number 1 - Default model)

	M.I.	Par Change
--	------	------------

Regression Weights: (Group number 1 - Default model)

	M.I.	Par Change
e21_1 <--- taahode	7.032	.364

۲.آ روش اجرای مدل‌سازی معادلات ساختاری در نرم‌افزار R برای مثال روان‌شناسی در فصل ۱

```
library("MVA")
set.seed(280875)
library("lattice")
lattice.options(default.theme = function() standard.theme("pdf", color = FALSE))
if (file.exists("deparse.R")) {
  if (!file.exists("figs")) dir.create("figs")
  source("deparse.R")
  options(prompt = "R> ", continue = "+ ", width = 64,
    digits = 4, show.signif.stars = FALSE, useFancyQuotes = FALSE)
  options(SweaveHooks = list(onefig = function() {par(mfrow = c(1,1))},
    twofig = function() {par(mfrow = c(1,2))},
    figtwo = function() {par(mfrow = c(2,1))},
    threefig = function() {par(mfrow = c(1,3))},
    figthree = function() {par(mfrow = c(3,1))},
    fourfig = function() {par(mfrow = c(2,2))},
    sixfig = function() {par(mfrow = c(3,2))},
    nomar = function() par("mai" = c(0, 0, 0, 0))))
}
alienation<- matrix(c(11.839, 6.947, 6.819, 4.783, -3.834, -21.899,
  6.947, 9.364, 5.091, 5.028, -3.889, -18.831,
  6.819, 5.091,12.532, 7.495, -3.841, -21.748,
  4.783, 5.028, 7.495, 9.986, -3.625, -18.775,
  -3.834,-3.889,-3.841,-3.625, 9.60, 35.522,
  -21.899,-18.831,-21.748,-18.775,35.522,450.283),
  ncol = 6, byrow = TRUE)
rownames(alienation) <- colnames(alienation) <- c("Anomia67",
  "Powles67", "Anomia71", "Powles71", "Educ", "SEI")
a <- cov2cor(alienation)
ord<- order.dendrogram(as.dendrogram(hclust(dist(a))))
panel.corrgram<- function(x, y, z, subscripts, at, level = 0.9, label = FALSE, ...)
```

```

{
require("ellipse", quietly = TRUE)
x <- as.numeric(x)[subscripts]
y <- as.numeric(y)[subscripts]
z <- as.numeric(z)[subscripts]
zcol<- level.colors(z, at = at, col.regions = grey.colors, ...)
for (i in seq(along = z)) {
ell<- ellipse(z[i], level = level, npoints = 50,
scale = c(.2, .2), centre = c(x[i], y[i]))
panel.polygon(ell, col = zcol[i], border = zcol[i], ...)
}
if (label)
panel.text(x = x, y = y, lab = 100 * round(z, 2), cex = 0.8,
col = ifelse(z < 0, "white", "black"))
}
print(levelplot(a[ord, ord], at = do.breaks(c(-1.01, 1.01), 20),
xlab = NULL, ylab = NULL, colorkey = list(space = "top"),
scales = list(x = list(rot = 90)),
panel = panel.corrgram, label = TRUE))

alienation_model<- specify.model( file = "alienation_model.txt")
alienation_sem<- sem(alienation_model, alienation, 932)

```

که مدل زیر را در نظر می‌گیرد.

```

SES ->Educ, NA, 1
SES -> SEI, lambda1, NA
Alienation67 -> Anomia67, NA, 1
Alienation67 -> Powles67, lambda2, NA
Alienation71 -> Anomia71, NA, 1
Alienation71 -> Powles71, lambda3, NA
SES -> Alienation67, beta1, NA
SES -> Alienation71, beta2, NA
Alienation67 -> Alienation71, beta3, NA
Educ<->Educ, theta1, NA

```

روش اجرای مدل سازی معادلات ساختاری در نرم افزار R برای مثال روان شناسی در فصل ۱ ۱۳۳

```
SEI <-> SEI, theta2, NA
SES <-> SES, delta0, NA
Anomia67 <-> Anomia67, theta3, NA
Powles67 <-> Powles67, theta4, NA
Anomia71 <-> Anomia71, theta5, NA
Powles71 <-> Powles71, theta6, NA
Alienation67 <-> Alienation67, delta1, NA
Alienation71 <-> Alienation71, delta2, NA
```

برآورد پارامترها عبارتند از

```
summary(alienation_sem)
```

```
Model Chisquare= 71.532 Df = 6 Pr(>Chisq) = 1.9825e-13
```

```
Chisquare (null model) = 2131.5 Df = 15
```

```
Goodness-of-fit index = 0.97514
```

```
Adjusted goodness-of-fit index = 0.913
```

```
RMSEA index = 0.10831 90% CI: (0.086636, 0.1315)
```

```
Bentler-Bonnett NFI = 0.96644
```

```
Tucker-Lewis NNFI = 0.9226
```

```
Bentler CFI = 0.96904
```

```
SRMR = 0.021256
```

```
AIC = 101.53
```

```
AICc= 72.056
```

```
BIC = 174.09
```

```
CAIC = 24.508
```

```
Normalized Residuals
```

```
Min.1st Qu. Median Mean 3rd Qu. Max.
```

```
-1.26000 -0.20900 -0.00001 -0.01510 0.24400 1.33000
```

```
R-square for Endogenous Variables
```

```
Educ SEI Alienation67 Anomia67 Powles67 Alienation71
```

0.6942 0.4205 0.3210 0.6602 0.6594 0.5764

Anomia71 Powles71

0.7045 0.6372

Parameter Estimates

Estimate Std Error z value Pr(>|z|)

lambda1 5.33054 0.429754 12.4037 2.4959e-35 SEI <--- SES

lambda2 0.88882 0.041530 21.4020 1.2788e-101 Powles67 <--- Alienation67

lambda3 0.84892 0.039965 21.2417 3.9310e-100 Powles71 <--- Alienation71

beta1 -0.61362 0.056441 -10.8718 1.5701e-27 Alienation67 <--- SES

beta2 -0.1744 0.053897 -3.2370 1.2078e-03 Alienation71 <--- SES

beta3 0.70463 0.053532 13.1629 1.4346e-39 Alienation71 <--- Alienation67

theta1 2.93613 0.499371 5.8797 4.1112e-09 Educ<-->Educ

theta2 260.93180 18.238709 14.3065 1.9934e-46 SEI <--> SEI

delta0 6.66387 0.640557 10.4032 2.3963e-25 SES <--> SES

theta3 4.02303 0.343389 11.7157 1.0596e-31 Anomia67 <--> Anomia67

theta4 3.18936 0.271559 11.7446 7.5264e-32 Powles67 <--> Powles67

theta5 3.70313 0.373384 9.9178 3.4850e-23 Anomia71 <--> Anomia71

theta6 3.62335 0.292079 12.4054 2.4430e-35 Powles71 <--> Powles71

delta1 5.30687 0.472629 11.2284 2.9577e-29 Alienation67 <--> Alienation67

delta2 3.73998 0.387452 9.6528 4.7851e-22 Alienation71 <--> Alienation71

Iterations = 86

۳.آ روش اجرای مدل سازی معادلات ساختاری ریج در نرم افزار SAS برای مثال واقعی در فصل ۲

NOTE: Copyright (c) 2002-2003 by SAS Institute Inc., Cary, NC, USA.

NOTE: SAS (r) 9.1 (TS1M3)

Licensed to ICFAI BUSINESS SCHOOL, Site 0049647001.

روش اجرای مدل سازی معادلات ساختاری ریج در نرم افزار SAS برای مثال واقعی در
فصل ۲ ۱۳۵

NOTE: This session is executing on the XP_PRO platform.

NOTE: SAS 9.1.3 Service Pack 4

NOTE: SAS initialization used:

real time	1.42 seconds
cpu time	0.45 seconds

```
1  **This program is to perform factor analysis using tetrachoric correlation matrix;  
2  *a constant a is added to  $S_a=S+aI$ , where S is the correlation matrix;  
3  *just for parameter estimation, no SE is involved;  
4  data raw;  
5  infile 'd:\programs\Item-factor\EPQaf1.dat'; *The full name of the data needs to be specified;  
6  input v1-v21; *The number of variables needs to be specified;  
7  
8  options ls=80;  
9  options nodate nonumber;  
10 title;
```

ERROR: Physical file does not exist, d:\programs\Item-factor\EPQaf1.dat.

NOTE: The SAS System stopped processing this step because of errors.

WARNING: The data set WORK.RAW may be incomplete. When this step was stopped
there were 0 observations and 21 variables.

NOTE: DATA statement used (Total process time):

real time	0.34 seconds
cpu time	0.03 seconds

```
11 proc iml;
```

NOTE: IML Ready

```
12 reset noname;
```

```

13  *-----*;
14  **** The following subroutine is to perform vech(.) function;
15  *-----*;
16  start vech(A,Va);
17  l=0;
18  p=nrow(A);
19  pstar=p*(p+1)/2;
20  Va=j(pstar,1,0);
21  do i=1 to p by 1;
22      do j=i to p by 1;
23          l=l+1;
24          Va[l]=A[j,i];
25      end;
26  end;
27  finish;
NOTE: Module VECH defined.
28  *-----*;
29  * creates duplication matrix as defined in Magnus&Neudecker*;
30  *-----*;
31  start DP(p, dup);
32  Dup=j(p*p,p*(p+1)/2,0);
33  count=0;
34  do j=1 to p;
35      do i=j to p;
36          count=count+1;
37          if i=j then do;
38              Dup[(j-1)*p+j, count]=1;
39          end;
40          else do;
41              Dup[(j-1)*p+i, count]=1;
42              Dup[(i-1)*p+j, count]=1;
43          end;
44      end;
45  end;

```

روش اجرای مدل سازی معادلات ساختاری ریج در نرم افزار SAS برای مثال واقعی در
فصل ۲ ۱۳۷

```
46 finish;

NOTE: Module DP defined.

47 *-----;
48 *standard normal pdf
49 *-----;
50 start pdf1(z, f1);
51     pi=3.14159265358979;
52     pi2=2*pi;
53     spi2=sqrt(pi2);
54     f1=exp(-z*z/2)/spi2;
55 finish;

NOTE: Module PDF1 defined.

56 *-----;
57 *standard bivariate normal pdf
58 *-----;
59 start pdf2(z1,z2,rho, f2);
60     pi=3.14159265358979;
61     pi2=2*pi;
62     rho12=1-rho*rho;
63     tt1=z1*z1-2*rho*z1*z2+z2*z2;
64     tt2=2*rho12;
65     f2=exp(-tt1/tt2)/( pi2*sqrt(rho12) );
66 finish;

NOTE: Module PDF2 defined.

67 *-----;
68 *Fisher scoring algorithm for tetrachoric correlation coefficient;
69 *-----;
70 start Estrho(htau1, htau2,n, n00,n01,n10,n11, rho0, hrho, diverg);
71 drho=.1;
72 *delta=.0000000001;
73 *delta=.1;
74 do i=1 to 50 while (abs(drho)>.0001);
75     pi00=PROBBNRM(htau1,htau2,rho0);
76     pi01=PROBNORM(htau1)-pi00;
```

```

77     pi10=PROBNORM(htau2)-pi00;
78     pi11=1-pi00-pi01-pi10;
79
80     run pdf2(htau1,htau2,rho0, phi2);
81     dl=(n00/pi00-n01/pi01-n10/pi10+n11/pi11)*phi2;
82     Eddl=n*(1/pi00+1/pi01+1/pi10+1/pi11)*phi2*phi2;
83     drho=d1/Eddl;
84     rho0=rho0+drho;
85 end;
86 hrho=rho0;
87 diverg=0;
88 if i=50 then diverg=1;
89 finish;
NOTE: Module ESTRHO defined.
90 *-----;
91 *Asymptotic covariance matrix using GEE;
92 *-----;
93 start ascov(n,p,x,htau, hrho,ncontig, omega);
94 ps=p*(p-1)/2;
95 pps=p+ps;
96 Amat11=j(p,p,0);
96 !           Amat12=j(p,ps,0);
96 !           Amat21=j(ps,p,0);
96 !           Amat22=j(ps,ps,0);
97 Bmat=j(pps,pps,0);
98
99 hpi=j(p,1,0);
99 !           *storing the probability of observing xij=1;
100 piphi=j(ps,5,0);
100!           *storing the probability and density of cell jk;
101 do j=1 to p;
102     hpi[j]=1-probnorm(htau[j]);
103 end;
104 jk=0;

```


روش اجرای مدل سازی معادلات ساختاری ریج در نرم افزار SAS برای مثال واقعی در
فصل ۲ ۱۳۹

```
105 do k=1 to p-1;
106     htau_k=htau[k];
107     probk=probnorm(htau_k);
108     do j=k+1 to p;
109         jk=jk+1;
110         htau_j=htau[j];
111         hrho_jk=hrho[j,k];
112         run pdf2(htau_j,htau_k,hrho_jk, phi_jk);
113         pi00=PROBBNRM(htau_j,htau_k,hrho_jk);
114         pi01=probnorm(htau_j)-pi00;
115         pi10=probk-pi00;
116         pi11=1-pi00-pi01-pi10;
117         piphi[jk,]=pi00||pi01||pi10||pi11||phi_jk;
118     end;
118!         *j;
119 end;
119!     *k;
120 do i=1 to n;
121     xi=x[i,]`;
122     dli1=j(p,1,0);
122!             *with respect to estimating tau;
123     do j=1 to p;
124         dli1[j]=xi[j]-hpi[j];
125     end;
126
127     dli2=j(ps,1,0);
127!             *with respect to estimating rho;
128     jk=0;
129     do k=1 to p-1;
130         do j=k+1 to p;
131             jk=jk+1;
132             nijk11=xi[j]*xi[k];
133             nijk10=xi[j]*(1-xi[k]);
134             nijk01=(1-xi[j])*xi[k];
```

```
135         nijk00=(1-xi[j])*(1-xi[k]);
136         dli2[jk]=(nijk00/piphi[jk,1]-nijk01/piphi[jk,2]
137             -nijk10/piphi[jk,3]+nijk11/piphi[jk,4])*piphi[jk,5];
138     end;
139 end;
140     dli=dli1//dli2;
141     dlidli=dli*dli`;
142     Bmat=Bmat+dlidli;
143 end;
144
145 pdftau=j(p,1,0);
146 do j=1 to p;
147     htauj=htau[j];
148     run pdf1(htauj, fj1);
149     pdftau[j]=fj1;
150 end;
151 Amat11=diag(pdftau);
152
153 jk=0;
154 ttvec=j(p,p,0);
155 do k=1 to p-1;
156     htauk=htau[k];
156!         htauk2=htauk*htauk;
157     fk1=pdftau[k];
158     do j=k+1 to p;
159         htauj=htau[j];
159!         htauj2=htauj*htauj;
160         fj1=pdftau[j];
161         jk=jk+1;
162         n00=ncontig[jk,1];
163         n01=ncontig[jk,2];
164         n10=ncontig[jk,3];
165         n11=ncontig[jk,4];
166         pi00=piphi[jk,1];
```

روش اجرای مدل سازی معادلات ساختاری ریج در نرم افزار SAS برای مثال واقعی در
فصل ۲ ۱۴۱

```
166! pi002=pi00**2;
167 pi01=pi01[jk,2];
167! pi012=pi01**2;
168 pi10=pi01[jk,3];
168! pi102=pi10**2;
169 pi11=pi01[jk,4];
169! pi112=pi11**2;
170 phiijk=pi01[jk,5];
171 hrhojk=hrho[j,k];
172 hrhojk2=hrhojk**2;
173 hrhojk21=1-hrhojk2;
174 shrhojk21=sqrt(hrhojk21);
175 dphiijk=phiijk*(
176 hrhojk*hrhojk21+htauj*htauk*(1+hrhojk2)-hrhojk*(htauj2+htauk2)
177 )/(hrhojk21**2);
178 coef1=n00/pi00-n01/pi01-n10/pi10+n11/pi11;
178! *E(coef1)=0; *coef1=0;
178! tt
178! vec[j,k]=coef1;
179 coef2=n10/pi102+n11/pi112;
180 coef3=n01/pi012+n11/pi112;
181 coef4=n00/pi002+n01/pi012+coef2;
182 Amat22[jk,jk]=coef1*dphiijk-coef4*phiijk*phiijk;
183
184 dphitauj=phiijk*(hrhojk*htauk-htauj)/hrhojk21;
185 dphitauk=phiijk*(hrhojk*htauj-htauk)/hrhojk21;
186 zjk=(htauj-hrhojk*htauk)/shrhojk21;
187 zkj=(htauk-hrhojk*htauj)/shrhojk21;
188 probjk=probnorm(zjk);
189 probkj=probnorm(zkj);
190 Amat21[jk,j]=coef1*dphitauj-coef4*fj1*probkj*phiijk+coef3*fj1*phiijk;
191 Amat21[jk,k]=coef1*dphitauk-coef4*fk1*probjk*phiijk+coef2*fk1*phiijk;
192 end;
193 end;
```

```

194 *maxtt=max(abs(ttvec)); *print "max tt=" maxtt;
195 *ind1=1:21;*ind2=0:21;
196 *ttvec=ind2//(ind1`||ttvec);
197 *print ttvec [format=f6.3];
198 Amat=(Amat11||Amat12)/(Amat21||Amat22)/n;
199 Bmat=Bmat/n;
200 Amatin=inv(Amat);
201 Omega=Amatin*Bmat*Amatin`;
202 finish;
NOTE: Module ASCOV defined.
203 *-----*;
204 **** The following subroutine generate the asymptotic covariance matrix
205 of
206 vech(\hat\rho)=[1, hrho_21,hrho_31,...,hrho_p1, 1, ..., hrho_p(p-1)], 1}
207 using
208 Omega=cov{vech(\hat\rho_ij)}=cov{[hrho_21,hrho_31,...,hrho_p(p-1)]};
209 *-----*;
210 start gmat(p,omega_sig, gamma);
211 ps=p*(p-1)/2;
212 pps=p+ps;
213 gamma=j(pps,pps,0);
214 jk=0;
215 index=0;
216 do k=1 to p;
217     do j=k to p;
218         jk=jk+1;
219!             *count all;
219         if j>k then do;
219!             *only record those below the diagonal;
220                 index=index||jk;
221         end;
222     end;
223 end;
224 index=index[2:(ps+1)];

```

روش اجرای مدل سازی معادلات ساختاری ریج در نرم افزار SAS برای مثال واقعی در
فصل ۲ ۱۴۳

```
225 gamma[index,index]=omega_sig;
226 finish;
NOTE: Module GMAT defined.
227
228 *-----*;
229 **** The following subroutine is to evaluate sigma0 and (dsigma0/dtheta)
230 in a correlation structure;
231 *-----*;
232
233 start sigdsig(a, p, theta0,vsig,Sigma,vdsig);
234 ps1=p*(p+1)/2;
235 Lamb=theta0[1:p];
236 ***** The following is to evaluate the function sigma;
237 LL=Lamb*Lamb`;
238 Psi=(a+1)*I(p)-diag(LL);
239 Sigma=LL+Psi;
240 run vech(Sigma,vsig);
241
242 ***** The following is to evaluate (dsigma/dtheta);
243
244 ***** DL is the derivative with Lambda;
245
246 DL=j(ps1,p,0);
247 do j=1 to p;
248     tt=j(p,1,0);
249     tt[j]=1;
250     tt1=j(p,p,0);
251     tt1[j,j]=-2*lamb[j];
252     ttt=tt*Lamb`+Lamb*tt`+tt1;
253     run vech(ttt,tttt);
254     DL[,j]=tttt;
255 end;
256 vdsig=DL;
257 finish;
```

NOTE: Module SIGDSIG defined.

```

258
259 ***-----***;
260 *Fisher scoring method for factor analysis by minimizing
261 F_ML=tr(S*Sig{-1})-log|S*Sig{-1}|-p;
262 ***-----***;
263 start minFml(a,p,theta0,Dup,sbig,vsig,fIRLS, weight, dswe, stdi,vdsig0,
263! diverg);
264 ep=.00000001;
265 *Fisher scoring begins here;
266 delta=.01;
267 do i=1 to 100 while (delta>ep);
268     run sigdsig(a,p, theta0,vsig0,Sigma0,vdsig0);
269     sign=inv(Sigma0);
270     ** weight given by normal theory ;
271     weight=0.5*dup`*(sign@sign)*dup;
272     dswe=vdsig0`*weight;
273     dwd=dswe*vdsig0;
274     stdi=inv(dwd);
275     eresdu=vsig-vsig0;
276     dtheta=stdi*dswe*eresdu;
277     theta0=theta0+0.5*dtheta;
278     delta=ssq(dtheta);
279 end;
280 if i<100 then do;
281     diverg=0;
282 **Test start here;
283     fIRLS=eresdu`*weight*eresdu;
284 end;
285 else do;
286     diverg=1;
287 end;
288 finish;

```

NOTE: Module MINFML defined.

روش اجرای مدل سازی معادلات ساختاری ریج در نرم افزار SAS برای مثال واقعی در
فصل ۲ ۱۴۵

```
289
290 *-----;
291 *-----;
292 **main program;
293 use raw;
WARNING: Data set WORK.RAW is empty.
```

```
statement : USE at line 293 column 1
294 read all var _num_ into x;
WARNING: End of File reached.
```

```
statement : READ at line 294 column 1
295 n=nrow(x);
295!           p=ncol(x);
296 x=x-j(n,p,1);
ERROR: (execution) Invalid operand to operation.
```

```
operation : J at line 296 column 6
operands  : N, P, *LIT1136
```

```
N      1 row      1 col      (numeric)
      0
```

```
P      1 row      1 col      (numeric)
      0
```

```
*LIT1136      1 row      1 col      (numeric)
      1
```

```
statement : ASSIGN at line 296 column 1
297
```

```

298
299 ps=p*(p-1)/2;
299!           pps=p+ps;
299!           a=0.5;
299!           *p/n;
299!           print "n,a=" n a;
300 delta=0.1;
300!           *delta=0.5;
301 *print
301! "-----";
302 *print x;
303 *print
303! "-----";
304 nvec=j(p,1,0);
ERROR: (execution) Invalid operand to operation.

```

operation : J at line 304 column 7

operands : P, *LIT1142, *LIT1143

P 1 row 1 col (numeric)

0

*LIT1142 1 row 1 col (numeric)

1

*LIT1143 1 row 1 col (numeric)

0

statement : ASSIGN at line 304 column 1

304! *storing the observed 1's for variables;

305 htau=j(p,1,0);

روش اجرای مدل سازی معادلات ساختاری ریج در نرم افزار SAS برای مثال واقعی در
فصل ۲ ۱۴۷

ERROR: (execution) Invalid operand to operation.

operation : J at line 305 column 7

operands : P, *LIT1144, *LIT1145

P 1 row 1 col (numeric)

0

*LIT1144 1 row 1 col (numeric)

1

*LIT1145 1 row 1 col (numeric)

0

statement : ASSIGN at line 305 column 1

305! *storing the thresholds;

306 do j=1 to p;

307 nj=sum(x[,j]);

308 nvec[j]=nj;

309 htau[j]=-probit(nj/n);

310 end;

311 print "hat\tau=" htau;

ERROR: Matrix HTAU has not been set to a value.

statement : PRINT at line 311 column 1

312 sigcov=x`*(i(n)-j(n,n,1)/n)*x/n;

ERROR: (execution) Invalid operand to operation.

operation : I at line 312 column 13

operands : N

N 1 row 1 col (numeric)

0

statement : ASSIGN at line 312 column 1

313 D_s=vecdiag(sigcov);

ERROR: (execution) Matrix has not been set to a value.

operation : VECDIAG at line 313 column 12

operands : SIGCOV

SIGCOV 0 row 0 col (type ?, size 0)

statement : ASSIGN at line 313 column 1

314 DSE=diag(j(p,1,1)/sqrt(D_s));

ERROR: (execution) Invalid operand to operation.

operation : J at line 314 column 11

operands : P, *LIT1149, *LIT1150

P 1 row 1 col (numeric)

0

*LIT1149 1 row 1 col (numeric)

1

*LIT1150 1 row 1 col (numeric)

1

statement : ASSIGN at line 314 column 1

روش اجرای مدل سازی معادلات ساختاری ریج در نرم افزار SAS برای مثال واقعی در
فصل ۲ ۱۴۹

```
315 scorr=DSE*sigcov*DSE;
```

```
ERROR: (execution) Matrix has not been set to a value.
```

```
operation : * at line 315 column 10
```

```
operands : DSE, SIGCOV
```

```
DSE      0 row      0 col      (type ?, size 0)
```

```
SIGCOV   0 row      0 col      (type ?, size 0)
```

```
statement : ASSIGN at line 315 column 1
```

```
315!           *sample Pearson correlation;
```

```
316 sig=scorr;
```

```
ERROR: (execution) Matrix has not been set to a value.
```

```
operation : MOVE at line 316 column 1
```

```
operands : SCORR
```

```
SCORR     0 row      0 col      (type ?, size 0)
```

```
statement : ASSIGN at line 316 column 1
```

```
316!           *j(p,p,.5) as starting values;
```

```
317 t0=j(p,1,.5);
```

```
ERROR: (execution) Invalid operand to operation.
```

```
operation : J at line 317 column 5
```

```
operands : P, *LIT1151, *LIT1152
```

```
P         1 row      1 col      (numeric)
```

*LIT1151 1 row 1 col (numeric)

1

*LIT1152 1 row 1 col (numeric)

0.5

statement : ASSIGN at line 317 column 1

318 q=nrow(t0);

319 df=ps-q;

319! print "df=" df;

320 *print nvec;

321 hrho=i(p);

ERROR: (execution) Invalid operand to operation.

operation : I at line 321 column 7

operands : P

P 1 row 1 col (numeric)

0

statement : ASSIGN at line 321 column 1

322 ncontig=j(ps,4,0);

ERROR: (execution) Invalid operand to operation.

operation : J at line 322 column 10

operands : PS, *LIT1154, *LIT1155

PS 1 row 1 col (numeric)

0

روش اجرای مدل سازی معادلات ساختاری ریج در نرم افزار SAS برای مثال واقعی در
فصل ۲ ۱۵۱

*LIT1154 1 row 1 col (numeric)

4

*LIT1155 1 row 1 col (numeric)

0

statement : ASSIGN at line 322 column 1

322! *contain the observations the 4 cells

323 in a 2 by 2 contingency table;

324 jk=0;

325 rvec=j(ps,1,0);

ERROR: (execution) Invalid operand to operation.

operation : J at line 325 column 7

operands : PS, *LIT1157, *LIT1158

PS 1 row 1 col (numeric)

0

*LIT1157 1 row 1 col (numeric)

1

*LIT1158 1 row 1 col (numeric)

0

statement : ASSIGN at line 325 column 1

326 do k=1 to p-1;

327 xk=x[,k];

```

328     nk=nvec[k];
329     do j=k+1 to p;
330         jk=jk+1;
331         xj=x[,j];
332         nj=nvec[j];
333         njk11=xj`*xk;
334         if njk11=0 then do;
335             njk11=delta;
336             *print "j k11=" j k;
337         end;
338         njk01=nk-njk11;
339         if njk01=0 then do;
340             njk01=delta;
341             *print "j k01=" j k;
342         end;
343         njk10=nj-njk11;
344         if njk10=0 then do;
345             njk10=delta;
346             *print "j k10=" j k;
347         end;
348         njk00=n-njk11-njk01-njk10;
349         if njk00=0 then do;
350             njk00=delta;
351             *print "j k00=" j k;
352         end;
353
354         ncontig[jk,]=njk00||njk01||njk10||njk11;
355
356         rhojk=sig[j,k];
356!             *starting value;
357         ** Fisher-scoring algorithm estimating tetrachoric correlations;
358         run Estrho(htau[j], htau[k],n, njk00,njk01,njk10,njk11, rhojk,
358! hrhojk, diverg);
359         if diverg=0 then do;

```

روش اجرای مدل سازی معادلات ساختاری ریج در نرم افزار SAS برای مثال واقعی در
فصل ۲ ۱۵۳

```
360          hrho[j,k]=hrhojk;
361          hrho[k,j]=hrhojk;
362          end;
363          else do;
364              print "For j=" j;
365              print "and k=" k;
366              print "the Fisher-scoring for rhojk does not converge within 50
366! iterations";
367          end;
368      end;
368!          *j;
369  end;
369!          *k;
370 *print "ncontig=" ncontig;
371 *-----;
372 *dht=htau-tau;
373 run vech(hrho, vhrho);
ERROR: (execution) Invalid operand to operation.
```

operation : J at line 20 column 5
operands : PSTAR, *LIT1004, *LIT1005

```
PSTAR      1 row      1 col      (numeric)

0

*LIT1004   1 row      1 col      (numeric)

1

*LIT1005   1 row      1 col      (numeric)

0
```

statement : ASSIGN at line 20 column 1

traceback : module VECH at line 20 column 1

NOTE: Paused in module VECH.

374 *print "hat\vrho=" Vhrho;

375

376 *run vech(hrho-sig, vhs);

377 *print vhrho vhs;

378

379 *drho=det(hrho); *print "det(hat\rho)=" drho;

380 *erho=eigval(hrho); *print "eigenv(hat\rho)=" erho;

381

382

383 run ascov(n,p,x,htau, hrho,ncontig, omega);

ERROR: (execution) Invalid operand to operation.

operation : J at line 96 column 9

operands : P, P, *LIT1044

P 1 row 1 col (numeric)

0

P 1 row 1 col (numeric)

0

*LIT1044 1 row 1 col (numeric)

0

statement : ASSIGN at line 96 column 1

traceback : module ASCOV at line 96 column 1

روش اجرای مدل سازی معادلات ساختاری ریج در نرم افزار SAS برای مثال واقعی در
فصل ۲ ۱۵۵

NOTE: Paused in module ASCOV.

384 *omega contains the asymptotic variance-covariance of both htau and hrho;

385 omega_sig=omega[(p+1):pps,(p+1):pps];

ERROR: (execution) Matrix has not been set to a value.

operation : [at line 385 column 16

operands : OMEGA, _TEM1002, _TEM1003, PPS

OMEGA 0 row 0 col (type ?, size 0)

_TEM1002 1 row 2 cols (numeric)

1 0

_TEM1003 1 row 1 col (numeric)

1

PPS 1 row 1 col (numeric)

0

statement : ASSIGN at line 385 column 1

385! *contains the asymptotic cov of hrho;

386 run gmat(p,omega_sig, gamma);

ERROR: (execution) Invalid operand to operation.

operation : J at line 213 column 8

operands : PPS, PPS, *LIT1109

PPS 1 row 1 col (numeric)

0

PPS 1 row 1 col (numeric)

0

*LIT1109 1 row 1 col (numeric)

0

statement : ASSIGN at line 213 column 1

traceback : module GMAT at line 213 column 1

NOTE: Paused in module GMAT.

386! *gamma has zero elements corresponding to

387 the diagonal elements of hrho;

388

389 *print omega;

390 *Fisher scoring method for factor analysis;

391 run DP(p, dup);

ERROR: (execution) Invalid operand to operation.

operation : J at line 32 column 8

operands : _TEM1001, _TEM1004, *LIT1012

_TEM1001 1 row 1 col (numeric)

0

_TEM1004 1 row 1 col (numeric)

0

*LIT1012 1 row 1 col (numeric)

0

روش اجرای مدل سازی معادلات ساختاری ریج در نرم افزار SAS برای مثال واقعی در
فصل ۲ ۱۵۷

statement : ASSIGN at line 32 column 3

traceback : module DP at line 32 column 3

NOTE: Paused in module DP.

392 theta0=t0;

ERROR: (execution) Matrix has not been set to a value.

operation : MOVE at line 392 column 1

operands : T0

T0 0 row 0 col (type ?, size 0)

statement : ASSIGN at line 392 column 1

393 sbig=hrho+a*I(p);

ERROR: (execution) Invalid operand to operation.

operation : I at line 393 column 14

operands : P

P 1 row 1 col (numeric)

0

statement : ASSIGN at line 393 column 1

394 *print "sbig=" sbig;

395 run vech(sbig, vsig);

ERROR: Module VECH called again before exit from prior call.

statement : RUN at line 395 column 1

396 run minFml(a,p,theta0,Dup,sbig,vsig, fIRLS,weight, dswe, stdi,vdsig0,
396! diverg1);

ERROR: (execution) Matrix has not been set to a value.

operation : [at line 235 column 12

operands : THETA0, *LIT1118, P

THETA0 0 row 0 col (type ?, size 0)

*LIT1118 1 row 1 col (numeric)

1

P 1 row 1 col (numeric)

0

statement : ASSIGN at line 235 column 1

traceback : module SIGDSIG at line 235 column 1

 module MINFML at line 268 column 5

NOTE: Paused in module SIGDSIG.

```

397 if diverg1=0 then do;
398     Vmata=weight-dswe`*stdi*dswe;
399     cst=trace(Vmata*Gamma)/df;
399!                                     *print "cst=" cst;
400     cst1=trace((Vmata*Gamma)*(Vmata*Gamma))/trace(Vmata*Gamma);
401     cst2=(trace(Vmata*Gamma)**2/trace((Vmata*Gamma)*(Vmata*Gamma)));
402     *print "cst1=" cst1; *print "cst2=" cst2;
403 *-----IRLS;
404     TIa=(n-1)*fIRLS;
404!                                     pIa=1-probchi(TIa,df);
405     TIra1=TIa/cst;
406     pIra1=1-probchi(TIra1,df);
407     TIRa2=TIa/cst1;
408     PIra2=1-probchi(TIra2,cst2);
409     print "T_IRLS,p_IRLS; T_IRLSra1,p_IRLSra1; T_IRLSra2,p_IRLSra2";
410     tttI=(TIa||pIa||TIra1||PIra1||TIRa2||PIra2);
411     print tttI;
412

```

```
413 end;
414 else do;
415     print "The Fisher-scoring algorithm for estimating theta does not
415! converge within 100 iterations";
416 end;
```

ERROR: (execution) Matrix has not been set to a value.

operation := at line 397 column 11

operands : DIVERG1, *LIT1173

DIVERG1 0 row 0 col (type ?, size 0)

*LIT1173 1 row 1 col (numeric)

0

statement : IF at line 397 column 1

مراجع

- [۱] آرست م، (۱۳۹۵)، پایان نامه ارشد: ”مقایسه رفتار برخی برآوردگرهای انقباضی بریج در مدل رگرسیون چندگانه”، دانشکده علوم ریاضی، دانشگاه صنعتی شاهرود،
- [۲] بروزئی بیدگلی م، (۱۳۹۳)، پایان نامه ارشد: ”بررسی رفتار برآوردگر جک نایف ریج”، دانشکده علوم ریاضی، دانشگاه صنعتی شاهرود،
- [۳] حسن زاده بشتیان م، (۱۳۸۸)، پایان نامه ارشد: ”برآوردگرهای انقباضی در مدل رگرسیون ریج با خطاهای دارای توزیع بیضی گون”، دانشکده علوم ریاضی، دانشگاه فردوسی مشهد،
- [۴] روزبه م، (۱۳۹۰)، پایان نامه دکتری: ”برآورد در مدل های خطی جزئی”، دانشکده علوم ریاضی، دانشگاه فردوسی مشهد،
- [۵] سلاجقه ع و آرشی م (۱۳۹۶)، چهل و هشتمین کنفرانس ریاضی ایران، ”مدل سازی معادلات ساختاری ریج”، ص ۱۲۲، همدان
- [۶] سلاجقه ع. و آرشی م، (۱۳۹۶) ”تأثیر سرمایه اجتماعی و رضایت شغلی بر تعهد سازمانی کارکنان ادارات ورزش و جوانان استان اصفهان و ارائه مدل”، نشریه مدیریت منابع انسانی در ورزش دانشگاه صنعتی شاهرود،
- [۷] سلاجقه ع، آرشی م و رحمانی م (۱۳۹۶)، کنفرانس روش های مدرن در قیمت گذاری های بیمه ای و آمارهای صنعتی، ”مدل معادلات ساختاری جریمه ریج”
- [۸] قاسمی و، (۱۳۹۲)، ”مدل سازی معادله ساختاری در پژوهش های اجتماعی با کاربرد **Amos Graphics** جلد اول، چاپ دوم، انتشارات جامعه شناسان
- [۹] مونت گمری د، پک ا، (۲۰۱۲)، (ترجمه ابراهیم رضوی پاریزی) ”مقدمه ای بر تحلیل رگرسیون خطی” جلد اول، چاپ چهارم، انتشارات دانشگاه شهید باهنر کرمان
- [۱۰] نجاریان س، (۱۳۹۰)، پایان نامه ارشد: ”بررسی رفتار برآوردگر ریج در مدل های خطی منفرد”، دانشکده علوم ریاضی، دانشگاه صنعتی شاهرود،

- [۱۱] نیرومند ح.ع.، (۱۳۸۷)، ” تحلیل رگرسیون خطی ابزاری برای تحقیق ” جلد اول، چاپ اول، انتشارات دانشگاه فردوسی مشهد
- [12] Akdeniz F. and Kaciranlar S. (1995), ” On the almost unbiased generalized Liu estimator and unbiased estimation of the Bias and MSE ” **Comm. Statist. Theo Meth** , 24, 1789-1797.
- [13] Babakus, E., Ferguson, C. E., and Joreskog, K. G. (1987), ” The sensitivity of confirmatory maximum likelihood factor analysis to violations of measurement scale and distributional assumptions. ” **Journal of Marketing Research** , 24, 224-229.
- [14] Bollen, K.A. (1989), ” Structural equations with latent variables. ” , Wiley, New York
- [15] Bollen, K.A. (2002), ” Latent variables in psychology and the social sciences. ” **Annual Review of Psychology** , 53, 605-634.
- [16] Browne, M. W. (2001), ” An Overview of Analytic Rotation in Exploratory Factor Analysis. ” **Multivariate Behavioral Research** , 36, 111-150.
- [17] Dahl, D. B. (2009), ” Xtable: Export Tables to Latex or Html. ” **R Package Version** , 1-5.
- [18] DiStefano, C. (2002), ” The impact of categorization with confirmatory factor analysis. ” **Structural Equation Modeling** , 9, 327-346.
- [19] Dolan, C. V. (1994), ” Factor analysis of variables with 2, 3, 5 and 7 response categories: A comparison of categorical variable estimators using simulated data. ” **British Journal of Mathematical and Statistical Psychology** , 47, 309-326.
- [20] Eysenck, H. J. and Eysenck, S. B. G. (1975), ” Manual of the Eysenck personality questionnaire. San Diego: Education and Industrial Testing Service. ”
- [21] Fan, J. and Li. R (2001), ” Variable Selection via Nonconcave Penalized Likelihood and Its Oracle Properties. ” **Journal of the American Statistical Association** , 96(456), 1348-1360.
- [22] GALTON F., (1886), ” Regression towards mediocrity in hereditary stature. ” **The Journal of the Anthropological Institute of Great Britain and Ireland** , 15, 246-263.
- [23] Gong, Y. (1983), ” Manual of EPQ Chinese revised version. ” , Changsha, China: Hunan Medical Institute
- [24] Hershberger, S.L. (2003), ” The growth of structural equation modeling: 1994–2001. ” **Structural Equation Modeling** , 10, 35-46.

- [25] Hoerl, A. E. and Kennard, R. W. (1970), " Ridge regression biased estimation for nonorthogonal problems " **Technometrics** , 12, 69-89.
- [26] Holzinger, Karl, J and Swineford, F (1939), " A Study in Factor Analysis: The Stability of a Bi-Factor Solution. " , Supplementary Educational Monographs
- [27] Joreskog, K. G. (1994), " On the estimation of polychoric correlations and their asymptotic covariance matrix. " **Psychometrika** , 59, 381-390
- [28] Jacobucci,R.(2017), "Package Regsem", <https://cran.r-project.org/web/packages/regsem/index.html>
- [29] Jacobucci, R, Grimm, K. J and McArdle, J. J (2016), " Regularized Structural Equation Modeling. " **Structural Equation Modeling: A Multidisciplinary Journal** , 23(4), 555-566.
- [30] Lawley, D.N., Maxwell, A.E. (1971), " Factor analysis as a statistical method, 2nd ed " , Elsevier, New York
- [31] Lee, S.-Y., Poon, W.-Y., and Bentler, P. M. (1995), " A two-stage estimation of structural equation models with continuous and polytomous variables. " **British Journal of Mathematical and Statistical Psychology** , 48, 339-358
- [32] Lei P.-W. (2009), " Evaluating estimation methods for ordinal data in structural equation modeling. " **Quality and Quantity** , 43, 495-507.
- [33] Liu, K. J. (2003), " Using Liu type estimator to combat multicollinearity " **Comm. Statist. Theo. Meth** , 32, 1009-1020.
- [34] MacCallum, R.C., Austin, J.T. (2000), " Applications of structural equation modeling in psychological research. " **Annual Review of Psychology** , 51, 201-226.
- [35] Magnus, J. R., and Neudecker, H. (1999), " Matrix differential calculus with applications in statistics and econometrics. " **New York: Wiley**
- [36] McArdle, J. J. (2005), " The Development of the Ram Rules for Latent Variable Structural Equation Modeling. " **Contemporary Psychometrics: A Festschrift for Roderick P. McDonald. Erlbaum.** , 225-273.
- [37] McQuitty, S. (1997), " Effects of employing ridge regression in structural equation models. " **Structural Equation Modeling** , 4, 244-252.

-
- [38] Muthen, B., and Satorra, A. (1995), " Technical aspects of Muthens LISCOMP approach to estimation of latent variable relations with a comprehensive measurement model. " **Psychometrika** , 60, 489-503.
- [39] Rosseel, Y. (2012), " Lavaan: An R Package for Structural Equation Modeling. " **Journal of Statistical Software** , 48(2), 1-36.
- [40] Schwarz, G. (1978), " Estimating the Dimension of a Model. " **The Annals of Statistics 6(2). Institute of Mathematical Statistics:** , 6(2), 461-464.
- [41] Spearman, C. (1904), " General intelligence" objectively determined and measured. " **American Journal of Psychology** , 15, 201-293.
- [42] Steiger, J. H, and Lind, J. C (1980), " Statistically Based Tests for the Number of Common Factors. " In Annual Meeting of the Psychometric Society, Iowa City, Ia. 758.
- [43] Swindel B. F. (1976), " Good ridge estimators based on prior information. " **Comm. Statist. Theo. Meth** , 5, 1065-1075.
- [44] Thurstone, L. L. (1935), " The Vectors of Mind. Chicago, IL: University of Chicago Press. "
- [45] Thurstone, L.L. (1947), " Multiple Factor Analysis. " , University of Chicago Press, Chicago
- [46] Tibshirani R. (1996), " Regression Shrinkage and Selection via the Lasso. " **JRoyal of the Statistical Society.Series B(Methodological)** , 58, 267-288.
- [47] Yuan, K. H., Wu, R., and Bentler, P. M. (2011), " Ridge structural equation modelling with correlation matrices for ordinal and continuous data " **British Journal of Mathematical and Statistical Psychology** , 64(1), 107-133.
- [48] Yuan, K.-H., and Chan, W. (2008), " Structural equation modeling with near singular covariance matrices. " **Computational Statistics and Data Analysis** , 52, 4842-4858.
- [49] Zhang, C-H. (2010), " Nearly Unbiased Variable Selection Under Minimax Concave Penalty. " **The Annals of Statistics. Institute of Mathematical Statistics:** , 38(2), 894-942.
- [50] Zou, H, and Hastie, T. (2005), " Regularization and Variable Selection via the ElasticNet. " **Journal of the Royal Statistical Society: Series B (Statistical Methodology) Wiley Online Library:** , 67(2), 301-320.

[51] Zou, H, Hastie, T and Tibshirani, R. (2006), " Sparse Principal Component Analysis "

[52] Zou, H, Hastie, T, and Tibshirani, T. (2007), " On the Degrees of Freedom of the Lasso. "

The Annals of Statistics. Institute of Mathematical Statistics: , 35(5), 2173-2192.

واژه‌نامه انگلیسی به فارسی

2-Stage Least Squares	حداقل مربعات دو مرحله‌ای
Absolute Fit Indices	شاخص‌های برازش مطلق
Adjusted Goodness-of-Fit Index	شاخص نیکویی برازش اصلاح شده
Akaike Information Criterion	معیار اطلاع آکاییک
Analysis of Covariance Structures	تحلیل ساختارهای کواریانس
Analysis of Mean Structures	تحلیل ساختارهای میانگین
Analysis of Moment Structures	تحلیل ساختارهای گشتاوری
Baddness-of-Fit Index	شاخص بدی برازش
Base Model	مدل مبنا
Bayse Information Criterion	معیار اطلاع بیز
Bentlet-Bonnet Index	شاخص بنتلر-بونت
Binary	دو مقوله‌ای
Browne-Cudeck Criterion	معیار براون-کادک
Categorical Ordinal Variable	متغیر مقوله‌ای ترتیبی
Causal Modeling	مدل سازی علی
Chi-Square	کای-دو
Comparative Fit Indices	شاخص برازش تطبیقی
Conceptual Model	مدل مفهومی
Covariance/Mean Matrix	ماتریس کواریانس / میانگین
Critical Ratio	نسبت بحرانی
Data-to-Model Fit	برازش داده‌ها به مدل
Default Model	مدل مفروض
Expected Cross-Validation Index	شاخص اعتبار متقاطع مورد انتظار
Equal Matrices	ماتریس‌های برابر
First-Order Confirmatory Factor Model	مدل عاملی تاییدی مرتبه اول
Fit Indices	شاخص‌های برازش

Fitness Index of Chi-Square	شاخص برازش کای-دو
Generalized Least Square	حداقل مربعات تعمیم یافته
Goodness-of-Fit Indices	شاخص‌های نیکویی برازش
Hoelter Index	شاخص هلتر
Identity Matrix	ماتریس همانی
Implied/Reproduce Matrix	ماتریس بازتولید شده
Incremental Fit Index	شاخص برازش افزایشی
Independence Model	مدل استقلال
Interface	میانجی
Latent Variables	متغیرهای پنهان
Likelihood Ratio Chi-Square	کای-دو نسبت درست‌نمایی
Main Diagonal	قطر اصلی
Mardias Coefficient	ضریب مردیا
Maximum Likelihood	ماکزیمم درست‌نمایی
Measurement Model	مدل اندازه‌گیری
Measurement Validity	اعتبار اندازه‌گیری
Modified Expected Cross Validation Index	شاخص اعتبار متقاطع مورد انتظار اصلاح شده
Model Estimation	برآورد مدل
Model Fit Indices	شاخص‌های برازش مدل
Model Identification	تشخیص مدل
Model Modification	اصلاح مدل
Model Specification	تدوین مدل
Model Test	آزمون مدل
Modification Indices	شاخص‌های اصلاح
Multivariate Normality	نرمال بودن چندمتغیره
Non-Central parameter	پارامتر غیر مرکزی
Non-Normed Fit Index	شاخص برازش هنجار نشده
Non-Redundant Element	عناصر غیرزاید
Nonrecursive Relation	رابطه بازگشتی
Normed Chi-Square	کای-دو هنجار شده
Normed Fit Index	شاخص برازش هنجار شده
Null Models	مدل‌های صفر
Numeric	عددی
Observed Variables	متغیرهای مشاهده شده

Ordered-Categorical Variables	متغیرهای مقوله‌ای ترتیبی
Over-Identified	فرامشخص
Parsimonious Fit Indices	شاخص‌های برازش مقتصد
Parsimonious Goodness-of-Fit Index	شاخص نیکویی برازش مقتصد
Parsimonious Normed Fit Index	شاخص برازش هنجار شده مقتصد
Parsimony Ratio	نسبت مقتصد بودن
Perfect Fit	برازش کامل
Reciprocal / Bidirectional	دو سویه / دو جهته
Recursive Models	مدل‌های غیر بازگشتی
Recursive Relation	رابطه غیر بازگشتی
Relative Chi-Square	کای-دو نسبی
Relative Difference	تفاوت نسبی
Relative Fit Index	شاخص برازش نسبی
Reproduce Covariance Matrix	ماتریس کواریانس بازتولید شده
Residual Matrix	ماتریس باقیمانده
Root Mean Square Error of Approximation	ریشه دوم میانگین مربعات خطای برآورد
Root Mean Squared Residual	ریشه دوم میانگین مربعات باقیمانده
Satorra-Bentler Corrected Chi-Square	کای-دو اصلاح شده ستورا-بنتلر
Saturated Model	مدل اشباع شده
Second-Order Confirmatory Factor Model	مدل عاملی تاییدی مرتبه دوم
Specified Model	مدل تدوین شده
Square Matrix	ماتریس مربع
String	رشته‌ای
Structural Equation Modeling	مدل‌سازی معادلات ساختاری
Structural Model	مدل ساختاری
Symmetric Matrix	ماتریس متقارن
Tucker-Lewis Index	شاخص توکر-لویس
Under-Identified	فرو مشخص
Unidirectional	یک جهته
Uniform Distribution	توزیع یکنواخت
Unobserved Variable	متغیر مشاهده نشده
Unordered Categorical (Nominal)	مقوله‌ای غیر ترتیبی (اسمی)
Unweight Least Square	حداقل مربعات غیروزی
Variance-Covariance Model	مدل واریانس-کواریانس

ماتریس صفر..... Zero Matrix

واژه‌نامه فارسی به انگلیسی

Model Test	آزمون مدل
Model Modeification	اصلاح مدل
Model Estimation	برآورد مدل
Data-to-Model Fit	برآزش داده‌ها به مدل
Perfect Fit	برآزش کامل
Non-Central Parameter	پارامتر غیرمرکزی
Analysis of Covariance Structures	تحلیل ساختارهای کواریانس
Analysis of Moment Structures	تحلیل ساختارهای گشتاوری
Analysis of Mean Structures	تحلیل ساختارهای میانگین
Factor Analysis	تحلیل عاملی
Model Specification	تدوین مدل
Model Identification	تشخیص مدل
Uniform Distribution	توزیع یکنواخت
Generalized Least Squares	حداقل مربعات تعمیم یافته
2-stage Least Squares	حداقل مربعات دو مرحله‌ای
Unweight Least Squares	حداقل مربعات غیر وزنی
Maximum Likelihood	ماکزیمم درست‌نمایی
Reciprocal / Bidirectional	دو سویه / دو جهته
Binary	دو مقوله‌ای
Nonrecursive Relation	رابطه بازگشتی
Recursive Relation	رابطه غیر بازگشتی
String	رشته‌ای
Root Mean Squared Residual	ریشه دوم میانگین مربعات باقیمانده
Root Meean Squared Error of Approximation	ریشه دوم میانگین مربعات خطای برآورد
Underestimation	زیر برآورد
Expected Cross-Validation Index	شاخص اعتبار متقاطع مورد انتظار

Modified Expected Cross-Validation Index . . .	شاخص اعتبار متقاطع مورد انتظار اصلاح شده
Baddnes-of-fit Index	شاخص بدی برازش
Incremental Fit Index	شاخص برازش افزایشی
Fitness Index of Chi-Square	شاخص برازش کای-دو
Relative Fit Index	شاخص برازش نسبی
Normed Fit Index	شاخص برازش هنجار شده
Non-Normed Fit Index	شاخص برازش هنجار نشده
Parsimonious Normed Fit Index	شاخص برازش هنجار شده مقتصد
Bentler-Bonett Index	شاخص بنتلر-بونت
Tucker-Lewis Index	شاخص توکر-لویس
Stability Index	شاخص ثبات
Adjusted Goodness-of-Fit Index	شاخص نیکویی برازش اصلاح شده
Parsimonious Goodness-of-Fit Index	شاخص نیکویی برازش مقتصد
Modification Indices	شاخص‌های اصلاح
Fit Indices	شاخص‌های برازش
Comparative Fit Indices	شاخص‌های برازش تطبیقی
Model Fit Indices	شاخص‌های برازش مدل
Absolute Fit Indices	شاخص‌های برازش مطلق
Parsimonious Fit Indices	شاخص‌های برازش مقتصد
Goodness-of-Fit Indices	شاخص‌های نیکویی برازش
Hoelters' Index	شاخص هلتر
Mardias' Coefficient	ضریب مردیا
Numeric	عددی
Over-identified	فرامشخص
Underidentified	فرو مشخص
Main Diagonal	قطر اصلی
Chi-Square	کای-دو
Satoora-Bentler Corrected Chi-Square	کای-دو اصلاح شده ستورا-بنتلر
Likelihood Ratio Chi-Square	کای-دو نسبت درست‌نمایی
Relative Chi-Square	کای-دو نسبی
Normed Chi-Square	کای-دو هنجار شده
Implied/Reproduced Matrix	ماتریس بازتولید شده
Residual Matrix	ماتریس باقیمانده
Zero Matrix	ماتریس صفر

Reproduce Covariance Matrix	ماتریس کواریانس بازتولید شده
Covariance/Mean Matrix	ماتریس کواریانس / میانگین
Symmetric Matrix	ماتریس متقارن
Square Matrix	ماتریس مربع
Equal Matrices	ماتریس‌های برابر
Identity Matrix	ماتریس همانی
unobserve Variable	متغیر مشاهده نشده
Categorical Ordinal Variable	متغیر مقوله‌ای ترتیبی
latent Variable	متغیرهای پنهان
Observed Variable	متغیرهای مشاهده شده
Independence Model	مدل استقلال
Saturated Model	مدل اشباع شده
Measurement Model	مدل اندازه‌گیری
Specified Model	مدل تدوین شده
Structural Model	مدل ساختاری
Causal Modeling	مدل‌سازی علی
Structural Equation Modeling	مدل‌سازی معادلات ساختاری
First-Order Confirmatory Factor Model	مدل عاملی تاییدی مرتبه اول
Secend-Order Confirmatory Factor Model	مدل عاملی تاییدی مرتبه دوم
Base Model	مدل مبنا
Default Model	مدل مفروض
Conceptual Model	مدل مفهومی
Unidentified Model	مدل نامشخص
Variance-Covariance Model	مدل واریانس - کواریانس
Null Models	مدل‌های صفر
Recursive Models	مدل‌های غیر بازگشتی
Akaike Information Criterion	معیار اطلاع آکاییک
Bayse Information Criterion	معیار اطلاع بیز
Brown-Cudeck Criterion	معیار براون - کادک
Interface	میانجی
Multivariate normality	نرمال بودن چند متغیره
Critical Ratio	نسبت بحرانی
Parsimony Ratio	نسبت مقتصد بودن
Unidirectional	یک جهته

Abstract

In some studies, there are variables that can not be directly measured or observed, often referred to as latent variables. Structural equation modeling provides a direct method for modeling latent variables, which combines structural model and measurement model to formulate structural equations. When there is a strong correlation between predictor variables, it is said multicollinearity exists. In such cases, the least squares estimator is not practical and often shrinkage estimators are used instead. One of such, is the ridge estimator. Here we examined the ridge method for solving structural equation modeling in ordinal and continuous data. In this regard, a constant value is added to the diagonal element of the correlation matrix. Then we minimized the objective function in the structural equation model on the basis of this change in order to estimate the model coefficients. We evaluated this method by a real example. Finally, we also gave a systematic modeling of the structural equations in the form of an example and compared the ridge method with some other alternatives.

Keywords: Correlation matrix, Latent variable, Multicollinearity, Ridge estimator, Shrinkage estimator, Structural equation modeling (SEM).



Shahrood University of Technology

Faculty Of Mathematical Sciences

MSc Thesis in: Mathematical Statistics

Ridge Structural Equation Model

By: Ali Salajegheh

Supervisor

Dr. Mohammad Arashi

September 2017