

تکنیک تحلیل کانونی و کاربرد آن در مسائل اقتصادی

دکتر اسماعیل شیرازی^۱

چکیده

روش تحلیل همبستگی کانونی یکی از شاخه های آماری می باشد که در زمینه مسائل اقتصادی، اجتماعی و کشاورزی از اهمیت بالایی در جهت یافتن ارتباط بین دو دسته از متغیرهای چندگانه برخوردار می باشد. در این مقاله به بررسی و چگونگی کاربرد این روش در مسائل اقتصادی پرداخته و پس از معرفی این روش، با ارائه یک مثال کاربردی جزئیات مورد بحث را روشن خواهیم نمود.

کلمات کلیدی: تحلیل همبستگی کانونی، تحلیل خوشه‌ای، رشد اقتصادی، تورم

^۱ . دانشگاه گنبد کاووس، دانشکده علوم پایه و فنی و مهندسی، گروه آمار و ریاضی



مقدمه

روش تحلیل مولفه‌های اساسی و تحلیل کانونی از روشهای آماری هستند که غالباً برای بررسی گروهی از متغیرهای همبسته مرتبط با یک یا چند حوزه مانند شاخصهای وضعیت اقتصادی اجتماعی، رضایت شغلی، سلامت، اعتبار شخصی و وضعیت سیاسی به کار میروند. مهمترین کاربردهای روشهای مذکور را میتوان در تجزیه و تحلیل نماگرهای چندگانه، اندازه‌گیری و شناخت ساختارهای پیچیده، شاخص‌سازی و کاهش داده‌ها جستجو نمود. این روشها به خصوص در شرایطی که ابعاد داده‌ها و ترکیب ساختار آنها کاملاً مشخص نیست مفید هستند. سالها است که از روشهای مذکور در علوم مختلف استفاده‌های زیادی میشود. یکی از کاربردهای آن در حوزه اقتصاد شاخص‌سازی است. به عنوان مثال با استفاده از این روش میتوان برای بازار داراییها شاخصهای قیمتی مناسبی را استخراج نمود. بازار داراییها یکی از بخشهای مهم و موثر در اقتصاد است که تغییر و تحولات آن نقش اساسی در تحولات سایر بازارها دارد و حتی گاهی مبین تحول عوامل ساختاری اقتصاد است. ارتباط سیستمی بخش مالی و واقعی اقتصاد ایران را میتوان در دوره‌های مختلف و با تحلیل سریهای زمانی نماگرهایی مانند رشد اقتصادی، تورم، شاخص‌های قیمت سهام و مسکن، میزان تسهیلات اعطایی بانکیها به بخشهای مختلف اقتصادی و سایر شاخصهای مرتبط بررسی و نتایج حاصل از این ارتباط سیستمی را در عمل به صورت پدیده‌های اقتصادی مشاهده نمود. در بررسی بسیاری از رخدادهای اقتصادی، مشاهده میشود عوامل مختلف با ابعاد مختلف و در حوزه‌های متفاوت بر پیدایش یک پدیده مانند تورم یا بیکاری موثر است. به کارگیری همه این عوامل در بررسیهای اقتصادسنجی نه تنها باعث ایجاد مشکلات فنی مانند هم‌خطی میشود بلکه باعث پیچیدگی مدل و افزایش اشتباهات آماری نیز میشود. بنابراین استفاده از تکنیکهای مختلف مانند روش مولفه‌های اساسی راهی برای کاهش این خطاها و دسترسی به نتایج بهتر میباشد. با توجه به شرایط تورمی اقتصاد ایران در سالهای اخیر و افزایش نقش داراییهایی چون زمین، مسکن، سهام و سایر داراییها در مکانیزم انتقال قیمتها و افزایش انتظارات تورمی، تحلیل دقیق چگونگی این ارتباط جهت شناسایی کانالهای ارتباطی متغیرهای مذکور با یکدیگر، بیش از پیش ضروری است. این بررسیها به سیاستگذاران اقتصادی کمک میکند تا بتوانند با داشتن اطلاعات بهتر و بیشتر، بسته‌های سیاستی به‌هنگام و مفیدتری را ارائه کنند. در ادامه در بخش اول ابتدا به روش تحلیل همبستگی کانونی پرداخته میشود و پس از شرح ویژگیها، نکات و پیشینه کاربرد این روش و چگونگی اجرای آن، رهیافت تحلیل عوامل و چگونگی محاسبه عوامل و پیشینه کاربردی آن بازگو و در پایان این بخش شباهت‌ها و تفاوت‌های هر دو روش شرح داده میشود. در بخش دوم کاربرد روشهای مذکور در اقتصاد ایران به صورت یک نمونه عملی بازگو میشود و شاخص قیمت داراییها با استفاده از روش مولفه‌های اساسی استخراج میشود. چگونگی تاثیر شاخص قیمت داراییها به عنوان یک نماگر مستقل و موثر بر تورم در بخش سوم این تحقیق ارائه می‌شود. در بخش پایانی جمع‌بندی و نتیجه‌گیری کل مباحث مطرح شده در هر یک از بخش‌ها بازگو میشود.

تحلیل همبستگی کانونی

تکنیک همبستگی کانونی، یک روش آماری چند متغیره‌ی مفید برای ارزیابی روابط خطی میان گروه‌هایی از متغیرها می‌باشد. به ویژه می‌تواند نقش برجسته‌ای به عنوان ابزاری اکتشافی به هنگام ارتباط داشتن متغیرهای چندگانه با یک مقوله تحلیلی ایفا کند [۱۰]. برای



به دست آوردن رابطه بین دو مجموعه متغیر مستقل و وابسته از روش تحلیل همبستگی کانونی (متداولترین حالت مدل خطی) استفاده می‌شود. این تکنیک آماری می‌تواند دو مجموعه متغیر را به طور همزمان در نظر بگیرد. همبستگی کانونی با دو مجموعه از داده‌ها آغاز می‌شود که شامل بردارهایی از مشاهدات انجام شده بر کلیه متغیرها می‌باشد. هدف همبستگی کانونی، با ایجاد X به عنوان یک بردار p بعدی از متغیرهای پیشین (مستقل) و Y به عنوان یک بردار q بعدی از متغیرهای ملاک (ملاک وابسته)، برای دستیابی به یک ترکیب خطی از متغیرهای پیشین است که دارای حداکثر همبستگی با یک ترکیب خطی از متغیرهای ملاک می‌باشد. این عمل بدین صورت می‌گیرد که به آن‌ها وزن‌های مختلفی می‌دهیم تا حداکثر همبستگی ممکن بین آن‌ها به دست آید [۱۱]. به طور مثال تشکیل دو ترکیب خطی به صورت $U=aX$ و $V=bY$ و همچنین ساختن یک ماتریس همبستگی $(p+q) \times (p+q)$ بین متغیرهای X_1 تا X_p که در آن p پارامتر فیزیکی و q پارامتر شیمیایی به دست آمده است.

$$\begin{array}{c}
 X_1 \quad X_2 \quad \dots \quad X_p \quad \quad Y_1 \quad Y_2 \quad \dots \quad Y_q \\
 \left[\begin{array}{cc}
 \begin{array}{c} X_1 \\ X_2 \\ \vdots \\ X_p \end{array} & \begin{array}{c} p \times p \text{ matrix} \\ A \end{array} & \begin{array}{c} p \times q \text{ matrix} \\ C \end{array} \\
 \hline
 \begin{array}{c} Y_1 \\ Y_2 \\ \vdots \\ Y_q \end{array} & \begin{array}{c} q \times p \text{ matrix} \\ C' \end{array} & \begin{array}{c} q \times q \text{ matrix} \\ B \end{array}
 \end{array} \right]
 \end{array}$$

از این ماتریس، یک ماتریس $B^{-1}C'A^{-1}C$ را می‌توان محاسبه نمود و مقادیر ویژه آن را با حل معادله زیر محاسبه کرد:

$$(B^{-1}C'A^{-1}C - \lambda I)b = 0 \quad (1)$$

در این حالت، مقادیر ویژه $\lambda_1 > \lambda_2 > \lambda_3 > \dots > \lambda_p$ عبارت از مربع مقادیر همبستگی بین دسته متغیرهای کانونیک می‌باشند. بردارهای ویژه b_1 تا b_p ضرایب متغیرهای Y را برای دسته متغیرهای کانونیک را نتیجه می‌دهد. ضریب U_i متغیر کانونیک i ام برای متغیرهای X ، توسط معادله زیر به دست می‌آیند. (عبدل آبادی و همکاران، ۱۳۹۱)

$$a_i = A^{-1}Cb_i \quad (2)$$

در نهایت، همبستگی بین دو جفت بردار U_i و V_i حاصل شده مورد بررسی قرار گرفته و به صورت ضریب همبستگی با مقادیر بین +۱ و -۱ ارائه می‌گردد. در این پژوهش از نرم‌افزار Statistics 24 برای محاسبه ضریب همبستگی با روش CCA استفاده شده است.

تحلیل خوشه‌ای

آنالیز خوشه‌ای یک روش طبقه بندی به منظور قرار دادن موجودیت‌های مشابه در دسته‌ها مختلف می‌باشد. طبقه بندی داده‌ها در این روش بر اساس فاصله همبستگی بین آن‌ها انجام می‌گیرد و داده‌هایی که فاصله آن‌ها از هم کمتر باشد در یک گروه قرار می‌گیرند.



هدف اصلی این روش، ایجاد طبقات مختلف که پراکنش و تفرق درون گروهی آن‌ها کمتر از بین‌گروهی آن‌ها باشد. مراحل طبقه‌بندی پارامترهای کیفی آب با روش تحلیل خوشه‌ای شامل تهیه ماتریس از داده‌ها، تعیین بار عاملی هر پارامتر به روش تحلیل عاملی، تلفیق طبقه‌ها و در نهایت گروه‌بندی و ترسیم دندوگرام است. در ابتدا به تعداد پارامترهای مورد بررسی خوشه وجود دارد و در آخرین مرحله، داده‌ها در تعداد کمتری خوشه تقسیم می‌شوند. دو پارامتر کیفی آب در صورتی متعلق به یک گروه می‌باشند که مقدار اندازه‌گیری شده آن‌ها نزدیک به هم باشند. اصولاً در این روش، خوشه‌بندی تا سطح سوم پیش برده می‌شود تا نتایج مطلوبی حاصل گردد [۳].

مثال کاربردی

یکی از موضوعات مورد توجه و تاثیرگذار در مسائل اقتصادی بازار دارایی‌ها است. از لحاظ تئوری اگر بازار دارایی‌ها از نظر اطلاعات کارا عمل کند و افراد عقلایی عمل کنند، قیمت دارایی‌ها منعکس‌کننده اطلاعات موجود درباره وقایع مورد انتظار است (اسمیت، ۱۹۹۹). اما در عمل بازار دارایی‌ها کامل نیستند و به دست آوردن اطلاعات با هزینه همراه است. همچنین بررسی بازار دارایی‌ها از آن نظر اهمیت می‌یابد که در شرایط معمولی، تغییر روند قیمت دارایی‌ها منعکس‌کننده تحول در عوامل ساختاری اقتصاد می‌باشد. اما در صورتیکه روند تغییر قیمت دارایی‌ها به شکل پایدار از روند طبیعی آن انحراف یابد حباب قیمت دارایی‌ها رخ می‌دهد. چنانچه این شرایط اقتصادی به درستی شناسایی نگردد پتانسیل بروز یک بحران مالی شکل خواهد گرفت. مانند آنچه در دهه ۹۰ در آمریکا به دنبال افزایش ناگهانی قیمت املاک رخ داد. دارایی‌ها را میتوان به دو دسته دارایی‌های مالی و واقعی تقسیم نمود. میتوان زمین و مسکن را جزء دارایی‌های واقعی و دارایی‌هایی نظیر طلا، ارز و سهام را بخشی از دارایی‌های مالی در نظر گرفت. عدم تعادل در هر یک از بازارها به راحتی به سایر بازارها و کل اقتصاد منتقل میشود. سرعت و میزان انتقال نه تنها به سهم دارایی‌ها در ثروت بخش خصوصی بلکه به سطح توسعه اقتصاد، به خصوص بازار مالی بستگی دارد. هرچند هنگامیکه روابط درونی میان دارایی‌ها (ارز، طلا، اوراق مشارکت، مستغلات و سایر موارد) و میان آنها با متغیرهای واقعی زیاد میشود، فرآیند انتقال مبهم و پیچیده میشود. به طور کلی برای اینکه بتوان رفتار و چگونگی این تعامل را در بین بازارها و در سطح کل اقتصاد بررسی نمود، باید شاخصی را که نشان‌دهنده تغییرات قیمت بازار دارایی‌ها است استخراج نمود. شاخص‌سازی و کاهش ابعاد داده‌ها یکی از روشهای آماری کارا در جهت بررسی وضعیت ابعاد مختلف اقتصاد و تحلیل روند سریهای زمانی میباشد. با توجه به سطح اطلاعات در دسترس و نوع مساله، تکنیکهای آماری مختلفی برای این کار وجود دارد که در بخش قبلی به دو روش اصلی آن اشاره شد.

محاسبات آماری

اطلاعات جمع آوری شد مرتبط با داده‌های قیمت سهام (SHPI)، دلار (EXPI)، سکه بهار آزادی (QPI)، زمین (GPI) و شاخص اجاری مسکن (HOPI) استفاده شده است. جهت همگن بودن مجموعه داده‌ها، متغیرهای مذکور، فصلی و به شکل شاخص قیمت، به سال پایه ۱۳۷۶ برای دوره زمانی ۱۳۷۳ تا ۱۳۸۵ و در نرم‌افزار Stata ۸ به کار رفته است. اولین گام در بررسی ارتباط بین متغیرها، بررسی ماتریس همبستگی آنها است. جدول شماره ۲-۲ نشان میدهد همبستگی بالایی بین همه داده‌ها وجود دارد و تنها شاخص اجاری مسکن (HOPI) همبستگی نسبتاً کمی با متغیرها دارد. بنابراین میتوان این متغیر را حذف نموده و به جای آن از



دومین همایش ملی اقتصاد کلان ایران
۱۱ اسفند ۱۳۹۵

دومین همایش ملی اقتصاد کلان ایران
یازدهم اسفندماه ۱۳۹۵



وزارت علوم، تحقیقات و فناوری
دانشگاه شهید باهنر اهواز

متغیر هزینه یک متر بنای مسکونی استفاده نمود. جدول شماره (۲-۳) جدول ماتریس همبستگی مجموعه داده‌های جدید در دوره ۱۳۷۳ تا ۱۳۸۵ را نشان می‌دهد. نتایج آزمون بارتلت شماره ۲-۴ نشان می‌دهد فرضیه صفر (عدم وجود همبستگی) برای مجموعه داده مذکور رد می‌شود، یعنی یک ارتباط قوی بین متغیرها وجود دارد. با توجه به این درجه همبستگی میتوان با استفاده از روش تحلیل مولفه‌های اساسی با کاهش تعداد متغیرها یک شاخص ترکیبی به نام شاخص قیمت دارایی‌ها ارائه نمود.

جدول شماره ۲-۲- ماتریس همبستگی بین متغیرها

	QPI	EXPI	SHPI	HOPI	GPI
QPI	۱				
EXPI	۰/۶۹۱۷	۱			
SHPI	۰/۷۸۹۳	۰/۶۰۱	۱		
HOPI	۰/۲۹۷	۰/۲۸۲۶	۰/۲۵۷۴	۱	
GPI	۰/۹۴۶۲	۰/۶۵۶۵	۰/۹۲۱۶	۰/۲۹۵۳	۱

جدول شماره ۲-۳- ماتریس همبستگی بین متغیرهای جدید

	QPI	EXPI	SHPI	HPI	GPI
QPI	۱				
EXPI	۰/۷۲۸۱	۱			
SHPI	۰/۸۳۵۸	۰/۶۷۳۴	۱		
HPI	۰/۹۵۴۸	۰/۸۱۴۶	۰/۸۹۵۹	۱	
GPI	۰/۹۵۶۲	۰/۷۱۴۷	۰/۹۳۹۴	۰/۹۷۲۱	۱

جدول شماره ۲-۴- نتایج آزمون بارتلت

Bartlett's sphericity test:	
Chi-square (observed value)	۴۷۰/۳۰۴
Chi-square (critical value)	۱۸/۳۰۷
DF	۱۰
One-tailed p-value	< ۰/۰۰۰۱
Alpha	۰/۰۵



جدول شماره ۲-۵ نتایج حاصل از برآورد را در نرمافزار ۸ Stata نشان میدهد. چنانکه مشاهده میشود مولفه‌های استخراج شده برابر است با تعداد متغیرهای استفاده شده، یعنی ۵ بردار، و همه پراکندگی مجموعه داده‌ها در بررسی در نظر گرفته شده است. با رسم نمودار میتوان بهترین مولفه را مشخص نمود.

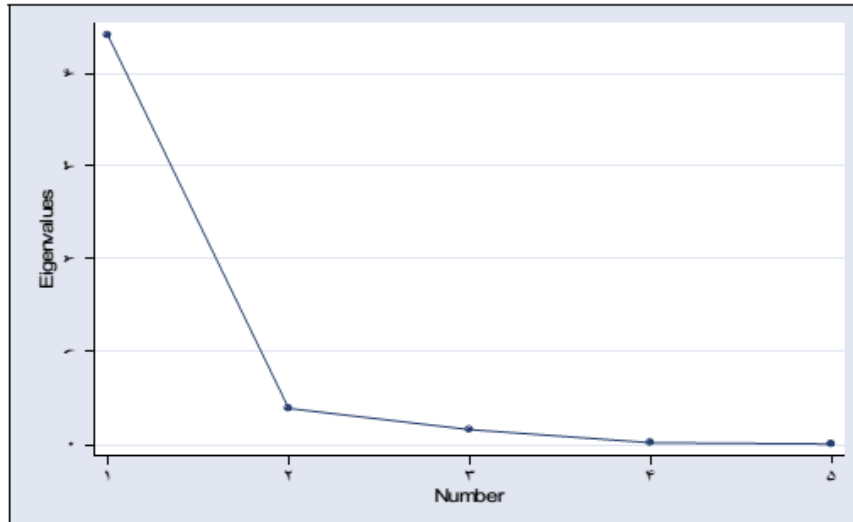
نمودار شماره ۲-۱ نشان میدهد مولفه اول بهترین انتخاب میباشد. چنانکه جدول شماره ۲-۵ نیز نشان میدهد مقدار ویژه اولین مولفه بزرگتر از یک میباشد و حدود ۸۹ درصد پراکندگی مجموعه داده‌ها توسط این مولفه بازگو میشود. به عبارت دیگر همه معیارها نشان دهنده این است که انتخاب مولفه اول کافی است. سهم هر یک از متغیرها در واریانس تبیین شده به وسیله هر عامل در جدول شماره ۲-۶ مشخص شده است. براین اساس مشاهده میشود سهم متغیرها در تبیین واریانس عامل اول نسبت به سایر ارقام مشابه مناسبتر میباشد. به عنوان مثال سهم هر یک از متغیرهای QPI, EXPI, SHPI, HPI و GPI در عامل F_1 به ترتیب عبارت است از: ۲۲/۲، ۱۹/۵، ۱۵/۶، ۲۰/۷ و ۲۱/۸۱.

جدول شماره ۲-۵- مولفه‌های اساسی

(principal components; ۵ components retained)				
Component	Eigenvalue	Difference	Proportion	Cumulative
۱	۴/۴۰۹۶۲	۴/۰۱۸۱۱	۰/۸۸۱۹	۰/۸۸۱۹
۲	۰/۳۹۱۵۱	۰/۲۲۶۲۸	۰/۰۷۸۳	۰/۹۶۰۲
۳	۰/۱۶۵۲۳	۰/۱۴۰۳۰	۰/۰۳۳۰	۰/۹۹۳۳
۴	۰/۰۲۴۹۳	۰/۰۱۶۲۲	۰/۰۰۵۰	۰/۹۹۸۳
۵	۰/۰۰۸۷۱	-	۰/۰۰۱۷	۱/۰۰۰۰

Eigenvectors					
Variable	۱	۲	۳	۴	۵
QPI	۰/۴۵۵۸۳	-۰/۱۴۷۷۹	-۰/۶۳۶۷۶	۰/۵۵۷۲۶	-۰/۲۳۳۱۶
EXPI	۰/۳۹۵۲۴	۰/۸۸۴۴۴	۰/۱۶۰۶۱	۰/۱۴۵۳۸	۰/۱۲۰۹۰
SHPI	۰/۴۴۲۳۴	-۰/۳۴۰۱۴	۰/۷۳۵۴۰	۰/۲۷۶۴۷	-۰/۲۶۷۲۰
HPI	۰/۴۷۱۴۵	۰/۰۰۱۱۲	-۰/۱۵۹۴۸	-۰/۷۴۹۹۱	-۰/۴۳۵۸۰
GPI	۰/۴۶۶۹۸	-۰/۲۸۳۲۴	-۰/۰۴۹۹۷	-۰/۱۷۱۷۹	۰/۸۱۸۳۵

نمودار شماره ۲-۱- مقادیر ویژه بردارها



جدول شماره ۲-۶- سهم متغیرها در تبیین واریانس عوامل

Contributions of the variables (%):					
	F ^۱	F ^۲	F ^۳	F ^۴	F ^۵
QPI	۲۰/۷۸۱	۲/۱۷۹	۴۰/۴۳۱	۳۱/۲۸۱	۵/۳۲۸
EXPI	۱۵/۶۱۸	۷۸/۲۷۱	۲/۵۹۲	۲/۰۷۴	۱/۴۴۴
SHPI	۱۹/۵۶۲	۱۱/۵۸۴	۵۴/۱۴۰	۷/۶۴۲	۷/۰۷۳
HPI	۲۲/۲۲۸	۰/۰۰۰	۲/۵۸۷	۵۵/۹۶۰	۱۹/۳۲۵
GPI	۲۱/۸۱۱	۷/۹۶۵	۰/۲۵۱	۳/۰۴۲	۶۶/۹۳۰

با توجه به ضرایب محاسباتی در بردار اول که در جدول شماره ۲-۵ ذکر شده است میتوان گفت بیشترین همبستگی را با مولفه اول به ترتیب متغیرهای مسکن، زمین، سکه، سهام و ارز دارا میباشند. بردار اول، که در خروجیهای کامپیوتری F_1 نامیده میشود را میتوان شاخص قیمت داراییها دانست و در مدل‌های مختلف اقتصادسنجی از آن استفاده نمود. با انتخاب این بردار، ترکیب خطی مولفه اول یا همان F_1 و متغیرهای اصلی با توجه به جدول ۲-۵ به صورت زیر میباشد:

$$F_1 = 0.45QPI + 0.39EXPI + 0.44SHPI + 0.47HOPI + 0.46GPI$$



ارتباط بین هر یک از متغیرها و مولفه‌ها را نیز میتوان با توجه به ضرایب محاسبه شده در جدول شماره ۲-۷ به صورت زیر نوشت:

$$QPI = 0.95F_1 - 0.92F_2 - 0.25F_3 - 0.88F_4 - 0.22F_5$$

$$EXPI = 0.83F_1 + 0.55F_2 + 0.65F_3 - 0.23F_4 + 0.11F_5$$

$$SHPI = 0.92F_1 - 0.21F_2 + 0.29F_3 - 0.44F_4 - 0.25F_5$$

$$HPI = 0.99F_1 - 0.65F_3 + 0.11F_4 - 0.41F_5$$

$$GPI = 0.98F_1 - 0.17F_2 - 0.2F_3 + 0.27F_4 + 0.76F_5$$

جدول شماره ۲-۷- ضرایب عوامل

Factor loadings:					
	F ^۱	F ^۲	F ^۳	F ^۴	F ^۵
QPI	۰/۹۵۷	-۰/۰۹۲	-۰/۲۵۹	-۰/۰۸۸	-۰/۰۲۲
EXPI	۰/۸۳۰	۰/۵۵۳	۰/۰۶۵	-۰/۰۲۳	۰/۰۱۱
SHPI	۰/۹۲۹	-۰/۲۱۳	۰/۲۹۹	-۰/۰۴۴	-۰/۰۲۵
HPI	۰/۹۹۰	۰/۰۰۰	-۰/۰۶۵	۰/۱۱۸	-۰/۰۴۱
GPI	۰/۹۸۱	-۰/۱۷۷	-۰/۰۲۰	۰/۰۲۷	۰/۰۷۶

بحث و نتیجه گیری

در این مطالعه با توجه به سطح دسترسی به اطلاعات و سهم عمده برخی از اقلام داراییها در سبد دارایی عاملان اقتصادی از شاخص قیمت مسکن، زمین، ارز، سکه و سهام و روش تحلیل مولفه‌های اساسی برای استخراج شاخص‌های قیمت دارایی‌ها استفاده شد. با استفاده از روش تحلیل کانونی در نرم‌افزار Stata پنج بردار دارایی استخراج شد که با توجه به سطح پوشش واریانس داده‌ها و سایر ملاک‌ها بردار اول برای تخمین شاخص کل قیمت دارایی‌ها انتخاب گردید.

مراجع

- ۱- مجموعه پژوهشهای بانک مرکزی/ مولف: فخری محدث/ سال ۱۳۸۹.
- ۲- فصلنامه‌های آماری بانک مرکزی جمهوری اسلامی ایران.
- ۳- تحلیل آماری چندمتغیری کاربردی/جانسون، ویچرن؛ ترجمه: حسینعلی نیرومند/دانشگاه فردوسی مشهد، ۱۳۷۸.