فصل ۷

مؤلفه های اصلی

۷ ـ ۱ مقدمه

در تحلیل چند متغیره، بزرگ بودن بعد بردار تصادفی، **X** ، اغلب در به دست آوردن روش‌های آماری مناسب برای نمونه تصادفی موجب مشکلاتی می‌گردد. حال می‌خواهیم با از دست دادن حداقل اطلاعات، بعد مشاهدات را تا حد قابل ملاحظه‌ای تقلیل دهیم.

این تفکر از آنجا 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 ثابت‌ها 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 بعد کمتر مورد مطالعه قرار دهیم.

روش مؤلفه‌های اصلی را ابتدا کارل پیرسن (۱۹۷۱) برای متغیرهای غیر آماری پیشنهاد کرد. در اکثر موارد یک تحلیل از مؤلفه‌های اصلی، ارتباط هایی که قبلا حدس زده شده را آشکار می سازد. تحلیل مؤلفه‌های اصلی در بیان‌های دیگر در مباحث رگرسیون چند متغیره، آنالیز گروه بندی و تجزیه عاملی نیز به کار گرفته می‌شود.

۷ ـ ۲ مؤلفه های اصلی بر اساس ماتریس کوواریانس جامعه

تحلیل مؤلفه‌های اصلی به ساختمان ماتریس کوواریانس به وسیله چند ترکیب خطی از متغیرهای اولیه، مربوط است. دو هدف عمده در اینجا پیگیری می‌شود.

۱- فشرده کردن داده،

۲- تفسیر اطلاعات.

با اینکه P مؤلفه‌ی اولیه در تغییر پذیری کل سیستم لازم است، اکثر اوقات این تغییر پذیری می‌تواند به وسیله تعداد کمتر، k،از مؤلفه‌های اصلی بیان شود.

فرض کنید بردار تصادفی $\mathbf{X} = (X_1, X_2, \dots, X_p)'$ دارای ماتریس کوواریانس Σ با مقادیر ویژه $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$ باشد. اگر ترکیبات خطی از $\mathbf{X}$ ، به فرم زیر را در نظر بگیریم

$$Y_i = \mathrm{l}_i' \mathbf{X} = l_{i1}X_1 + l_{i2}X_2 + \dots + l_{ip}X_p \qquad i = 1,2,\dots,p$$

واریانس Y_iها و کوواریانس Y_j و Y_k برابر است با

$$Var(Y_i) = \mathbf{I}_i' \Sigma\, \mathbf{I}_i \qquad i = 1,2,\dots,p$$

$$Cov(Y_i, Y_k) = \boldsymbol{I}_i' \Sigma\, \boldsymbol{I}_k \qquad i \neq k \qquad i,k = 1,2,\dots,p$$

برای ساختن مؤلفه های اصلی $\boldsymbol{I}_i$ *ها به نحوی انتخاب شوند که* $\boldsymbol{Y}_i$ *ها ناهمبسته با بیشترین واریانس باشند به عبارت دیگر*

اولین مؤلفه اصلی = ترکیب خطی $\boldsymbol{I}_1' \mathbf{X}$ *با بیشترین* $Var\left(\boldsymbol{I}_1' \boldsymbol{X}\right)$ *و فرض* $\mathbf{I}_1' \mathbf{I}_1 = 1$,

دومین مؤلفه اصلی = ترکیب خطی $\boldsymbol{I}_2' \mathbf{X}$ *با بیشترین* $Var\left(\boldsymbol{I}_2' \boldsymbol{X}\right)$ *و فرض* $\mathbf{I}_2' \mathbf{I}_2 = 1$ و

$$Cov\left(\boldsymbol{I}_1\boldsymbol{X}, \boldsymbol{I}_2\boldsymbol{X}\right) = 0$$

$$\vdots \qquad\qquad\qquad\qquad \vdots$$

i *دومین مؤلفه اصلی = ترکیب خطی* $\boldsymbol{I}_i' \mathbf{X}$ *با بیشترین* $Var(\boldsymbol{I}_i' \boldsymbol{X})$ *و فرض* $\mathbf{I}_i' \mathbf{I}_i = 1$و

برای $k < i$ $\qquad Cov\left(\boldsymbol{I}_1\boldsymbol{X}, \boldsymbol{I}_2\boldsymbol{X}\right) = 0$

$$\vdots \qquad\qquad\qquad\qquad \vdots$$

به کمک روش لاگرانژ، ثابت می گردد: اگر زوج های بردار ویژه- مقدار ویژه *ماتریس* Σ , $\left(\lambda_1, \boldsymbol{e}_1\right), \dots, \left(\lambda_p, \boldsymbol{e}_p\right)$ با $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$ در دست باشند آنگاه

$$\lambda_1 = \boldsymbol{e}_1' \,\Sigma\, \boldsymbol{e}_1 = \max_{\{\boldsymbol{I} \;:\; \boldsymbol{I}'\boldsymbol{I}=1\}} \boldsymbol{I}'\boldsymbol{\Sigma}\boldsymbol{I}$$

$$\lambda_2 = \boldsymbol{e}_2' \,\Sigma\, \boldsymbol{e}_2 = \max_{\{\boldsymbol{I} \;:\; \boldsymbol{I}'\boldsymbol{I}=1,\; \boldsymbol{I}'\boldsymbol{e}_1=\mathbf{0}\}} \boldsymbol{I}'\boldsymbol{\Sigma}\boldsymbol{I}$$

$$\lambda_3 = \boldsymbol{e}_3' \,\Sigma\, \boldsymbol{e}_3 = \max_{\{\boldsymbol{I} \;:\; \boldsymbol{I}'\boldsymbol{I}=1,\; \boldsymbol{I}'\boldsymbol{e}_1=\mathbf{0},\; \boldsymbol{I}'\boldsymbol{e}_2=0\}} \boldsymbol{I}'\boldsymbol{\Sigma}\boldsymbol{I}$$

$$\vdots$$

و از روی این جواب, i امین مؤلفه اصلی برابر است با

$$Y_i = \boldsymbol{e}_i' \boldsymbol{X} = e_{1i}\, X_1 + e_{2i}\, X_2 + \cdots + e_{pi}\, X_p \qquad i = 1,2,\dots,p$$

با این انتخاب داریم

$$Var(Y_i) = \boldsymbol{e}_i' \Sigma\, \boldsymbol{e}_i = \lambda_i \qquad i = 1,2,\dots,p$$

$$Cov(Y_i, Y_k) = \boldsymbol{e}_i' \Sigma\, \boldsymbol{e}_k = ۰ \qquad\qquad i \neq k$$

در صورت برابری برخی λ_i ها، $\boldsymbol{e}_i$ ها یکی نخواهند شد و در نتیجه Y_i های متناظر نیز یکی نیستند. در واقع با داشتن دو ماتریس

$$\boldsymbol{\Lambda} = diag(\lambda_۱, \ldots, \lambda_p)$$

و

$$\boldsymbol{P} = [\boldsymbol{e}_۱, \ldots, \boldsymbol{e}_p]$$

رابطه زیر برقرار است

$$\Sigma = \boldsymbol{p}\, \boldsymbol{\Lambda}\, \boldsymbol{p}'.$$

مهمترین ویژگی که دراین تبدیل مشاهده می‌شود رابطه بین مجموع مقادیر ویژه و مجموع واریانس های مؤلفه های $\boldsymbol{X}$ می باشد.

$$\sum_{i=۱}^{p} Var(X_i) = tr(\Sigma)$$

$$= tr(\boldsymbol{\Lambda})$$

$$= \sum_{i=۱}^{p} Var(Y_i)$$

و یا

$$\text{واریانس کل در جامعه} = \sigma_{۱۱} + \sigma_{۲۲} + \cdots + \sigma_{pp} = \lambda_۱ + \lambda_۲ + \cdots + \lambda_p$$

به طوری که دیده می‌شود در این مجموع، متغیرهایی که واریانس کوچکتری دارند نقش کمی در این حاصل پیدا کرده و از این رو را می‌توان نادیده گرفت. پس متغیر ها به تعداد کمتری، $k \leq p$ آنها، کاهش می‌یابند.

اختلاف اطلاعات بین کل واریانس $tr\Sigma$ و مجموع واریانس K مقدار گیریم، می نظر در اصلی مولفه K را باید بنحوی انتخاب کنیم که این اختلاف عدد کوچکی باشد.

$$tr\Sigma - var\left(\sum_{i=۱}^{k} Y_i\right) = \sum_{i=۱}^{p} \lambda_i - \sum_{i=۱}^{k} \lambda_i$$

با داشتن k مولفه اصلی، یک سوال طبیعی مطرح است: آیا گروهی از متغیر ها مانند $Z = T'\mathbf{X}$ (که در آن T یک ماتریس $k \times p$ است) وجود دارد که مجموع واریانس متغیرهای آن بیشتر از مجموع واریانس K مولفه اصلی شود؟ پاسخ منفی است زیرا عبارت

$$tr\Sigma - tr\big(cov(T'\boldsymbol{X})\big) = tr\Sigma - tr(T'\Sigma T)$$

به ازای $T = [\boldsymbol{e}_۱, \ldots, \boldsymbol{e}_k]$ کمترین مقدار می‌شود.

نقش تاثیر i امین مولفه اصلی با کمیت زیر قابل اندازه گیری است:

$$\frac{\lambda_i}{\lambda_۱ + \lambda_۲ + \cdots + \lambda_p}$$

اگر مجموع نسبت فوق تا مولفه اصلی kام،

$$\frac{\lambda_۱ + \lambda_۲ + \cdots + \lambda_k}{\lambda_۱ + \lambda_۲ + \cdots + \lambda_p}$$

حداکثر بین ۰/۸ تا ۰/۹ باشد می توان مسئله را با K مولفه های اصلی به جای p مولفه انجام داد.

از نقطه نظر دیگر، با فرض اینکه بردار میانگین جامعه برابر صفر باشد، می توان کوواریانس بین مولفه های اصلی و مولفه های اولیه را محاسبه کرد و تاثیر هر یک از مولفه های $\mathbf{X}$ را در ساختار مولفه های اصلی مورد ارزیابی قرارداد. به عبارت دیگر داریم:

$$Cov(Y_i, X_k) = E(Y_i X_k) - E(Y_i)E(X_k)$$
$$= E(Y_i X_k)$$
$$= E(e_i'\boldsymbol{X}\, X_k)$$
$$= e_i' E(\boldsymbol{X}\, X_k)$$
$$= e_i' E((X_۱ X_k, \ldots, X_p X_k)')$$
$$= e_i'(\sigma_{k۱}, \ldots, \sigma_{kp})'$$
$$= \lambda_i e_{ki} \qquad (\Sigma e_i = \lambda_i e_i)$$

وضرب همبستگی بین مولفه های اصلی و مولفه های اولیه

$$\rho_{Y_i, X_k} = \frac{e_{ki}\lambda_i}{\sqrt{\lambda_i}\sqrt{\sigma_{kk}}}$$
$$= \frac{e_{ki}\sqrt{\lambda_i}}{\sqrt{\sigma_{kk}}} \qquad i, k = ۱,۲, \ldots, p \qquad (۱ - ۲ - ۷)$$

ضرایب در ساختار مولفه های اصلی ضریبی از همبستگی خود با مولفه های اولیه می باشد. در جدول زیر همبستگی بین مولفه‌های اصلی و مولفه‌های اولیه تنظیم شده است.

مولفه های اولیه	مولفه های اصلی			
	Y_1	Y_2	$\dots$	Y_p
X_1	$\lambda_1^{\frac{1}{2}} e_{11} \sigma_{11}^{-\frac{1}{2}}$	$\lambda_2^{\frac{1}{2}} e_{12} \sigma_{11}^{-\frac{1}{2}}$	$\dots$	$\lambda_p^{\frac{1}{2}} e_{1p} \sigma_{11}^{-\frac{1}{2}}$
X_2	$\lambda_1^{\frac{1}{2}} e_{21} \sigma_{22}^{-\frac{1}{2}}$	$\lambda_2^{\frac{1}{2}} e_{22} \sigma_{22}^{-\frac{1}{2}}$	$\dots$	$\lambda_p^{\frac{1}{2}} e_{2p} \sigma_{22}^{-\frac{1}{2}}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
X_p	$\lambda_1^{\frac{1}{2}} e_{p1} \sigma_{pp}^{-\frac{1}{2}}$	$\lambda_2^{\frac{1}{2}} e_{p2} \sigma_{pp}^{-\frac{1}{2}}$	$\dots$	$\lambda_p^{\frac{1}{2}} e_{pp} \sigma_{pp}^{-\frac{1}{2}}$

مثال ۷ – ۲ -۱: فرض کنید متغیرهای تصادفی X_1, X_2, X_3 دارای ماتریس کوواریانس زیر هستند. با صفر گرفتن برخی درایه های ماتریس کوواریانس، تاثیر ناهمبستگی متغیرها را در مولفه‌های اصلی بررسی می‌کنیم.

$$\Sigma = \begin{bmatrix} 1 & -2 & 0 \\ -2 & 5 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

مقادیر ویژه و بردارهای ویژه Σ عبارتند از:

$$\lambda_1 = 583 \quad , \quad \boldsymbol{e}_1' = [0.383, \quad -0.924, \quad 0.000]$$

$$\lambda_2 = 2.00 \quad , \quad \boldsymbol{e}_2' = [0.000, \quad 0.000, \quad 1.000]$$

$$\lambda_3 = 0.17 \quad , \quad \boldsymbol{e}_3' = [0.924, \quad -0.383, \quad 0.000]$$

و مولفه‌های اصلی

$$Y_1 = \boldsymbol{e}_1' \boldsymbol{X} = 0.383X_1 - 0.924X_2$$

$$Y_{۲} = \boldsymbol{e}'_{۲}\,\boldsymbol{X} = X_{۳}$$

$$Y_{۳} = \boldsymbol{e}'_{۳}\,\boldsymbol{X} = ۰.۹۲۴X_{۱} + ۰.۳۸۳X_{۲}$$

ملاحظه می‌شود که $\frac{\lambda_{۱}}{\lambda_{۱}+\lambda_{۲}+\lambda_{۳}} = ۰.۷۳$ و $\frac{\lambda_{۱}+\lambda_{۲}}{\lambda_{۱}+\lambda_{۲}+\lambda_{۳}} = ۰.۹۸$ یعنی با به کار گیری $Y_{۱}$ و $Y_{۲}$ به جای سه متغیر اولیه اطلاعات ناچیزی را از دست می‌دهیم.

$$\rho_{Y_{۱},X_{۱}} = e_{۱۱}\sqrt{\frac{\lambda_{۱}}{\sigma_{۱۱}}} = \frac{۰.۳۸۳\sqrt{۵.۸۳}}{\sqrt{۱}} = ۰.۹۲۵$$

$$\rho_{Y_{۱},X_{۲}} = e_{۲۱}\sqrt{\frac{\lambda_{۱}}{\sigma_{۲۲}}} = \frac{-۰.۹۲۴\sqrt{۵.۸۳}}{\sqrt{۵}} = -۰.۹۹۸$$

با این دو مقدار، نقش و تاثیر $X_{۱}$ و $X_{۲}$ در ساختار $Y_{۱}$ نشان داده می‌شود. همچنین

$$\rho_{Y_{۲},X_{۳}} = e_{۲۳}\sqrt{\frac{\lambda_{۲}}{\sigma_{۳۳}}} = \sqrt{\frac{۲}{۲}} = ۱$$

$$\rho_{Y_{۲},X_{۱}} = \rho_{Y_{۲},X_{۲}} = ۰$$

با توجه به $Y_{۲}$ این مقادیر قابل پیش‌بینی بود.

همبستگی چند گانه بین i امین مولفه اولیه و K تا مولفه اصلی اول را می‌توان محاسبه کرد:

$$\rho^{۲}{}_{X_i,Y_{۱},\dots,Y_k} = \sum_{j=۱}^{k} \lambda_i e_{ij}^{۲} / \sigma_{ii}.$$

۳-۷ بدست آوردن مولفه‌های اصلی از متغیرهای استاندارد

مؤلفه های اصلی را می توان از متغیرهای استاندارد شده ی زیر نیز بدست آورد.

$$Z_i = \frac{(X_i - \mu_{i۱})}{\sqrt{\sigma_{ii}}} \qquad\qquad i = ۱,۲,\dots,p$$

در نمایش ماتریسی اگر $D_\sigma = diag(\sigma_{۱۱}, \dots, \sigma_{pp})$ و $\boldsymbol{Z}' = (Z_{۱}, \dots, Z_p)$ باشد آنگاه

$$\boldsymbol{Z} = (D_\sigma^{\frac{۱}{۲}})^{-۱}\,(\mathbf{X} - \boldsymbol{\mu})$$

واضح است که $E(\boldsymbol{Z}) = \boldsymbol{o}$ و

$$Cov(\boldsymbol{Z}) = (D_\sigma^{\frac{۱}{۲}})^{-۱}\, \Sigma (D_\sigma^{\frac{۱}{۲}})^{-۱} = \boldsymbol{\rho}$$

مولفه‌های اصلی را از روی ماتریس همبستگی، $\boldsymbol{\rho}$ ، به دست می‌آوریم. مجددا Y_i ها مولفه‌های اصلی، λ_i مقادیر ویژه ماتریس $\boldsymbol{\rho}$ و $\boldsymbol{e}_i$ ها بردار ویژه ماتریس $\boldsymbol{\rho}$ هستند

$$Y_i = \boldsymbol{e}_i' \, \boldsymbol{Z} = \boldsymbol{e}_i' \, (D_\sigma^{\frac{۱}{۲}})^{-۱} \, (\mathbf{X} - \boldsymbol{\mu}) \qquad i = ۱,۲, \dots, p$$

بعلاوه

$$\sum_{i=۱}^{p} Var\,(Y_i) = \sum_{i=۱}^{p} Var(Z_i) = p$$

و

$$\rho_{Y_i, Z_k} = e_{ki}\sqrt{\lambda_i} \qquad i,k = ۱,۲, \dots, p \qquad (۱\text{-}۳\text{-}۷)$$

k مولفه اصلی مورد بحث تا رسیدن نسبت $\sum_{i=۱}^{k} \frac{\lambda_i}{p}$ به ۰/۸ تعیین می‌گردد.

مثال۷-۳-۱: بنابر ماتریس کوواریانس

$$\Sigma = \begin{bmatrix} ۱ & ٤ \\ ٤ & ۱۰۰ \end{bmatrix}$$

ماتریس همبستگی حاصله برابر است با

$$\begin{bmatrix} ۱ & ۰.٤ \\ ۰.٤ & ۱ \end{bmatrix}$$

بردارها و مقادیر ویژه ماتریس Σ عبارتند از

$$\lambda_۱ = ۱۰۰.۱٦ \quad , \qquad \boldsymbol{e}_۱' = [۰.۰٤۰, \quad ۰.۹۹۹]$$

$$\lambda_۲ = ۰.۸٤ \quad , \qquad \boldsymbol{e}_۲' = [۰.۹۹۹, \quad -۰.۰٤۰]$$

و به طور مشابه مقادیر و بردارهای ویژه ماتریس $\boldsymbol{\rho}$ عبارت است از:

$$\lambda_۱ = ۱ + \rho = ۱.٤ \quad , \qquad \boldsymbol{e}_۱' = [۰.۷۰۷, \quad ۰.۷۰۷]$$

$$\lambda_۲ = ۱ - \rho = ۰.٦ \quad , \qquad \boldsymbol{e}_۲' = [۰.۷۰۷, \quad -۰.۷۰۷]$$

مولفه های اصلی براساس ماتریس کوواریانس جامعه و ماتریس همبستگی جامعه عبارتند از

$$\Sigma : \begin{cases} Y_۱ = ۰.۰٤۰X_۱ + ۰.۹۹۹X_۲ \\ Y_۲ = ۰.۹۹۹X_۱ - ۰.۰٤۰X_۲ \end{cases}$$

و

$$\boldsymbol{\rho}:\begin{cases} Y_۱ = ۰.۷۰۷Z_۱ + ۰.۷۰۷Z_۲ \\ \quad = ۰.۷۰۷\left(X_۱-\mu_۱\right) + ۰.۷۰۷(X_۲-\mu_۲) \\ Y_۲ = ۰.۷۰۷Z_۱ - ۰.۷۰۷Z_۲ \\ \quad = ۰.۷۰۷\left(X_۱-\mu_۱\right) - ۰.۷۰۷\left(X_۲-\mu_۲\right) \end{cases}$$

بزرگ بودن واریانس، نقش $X_۲$ را در مولفه اصلی اول از Σ مهم جلوه می‌دهد. بعلاوه واریانس نسبی مولفه اول $\frac{\lambda_۱}{\lambda_۱+\lambda_۲} = ۰.۹۹۲$ از کل واریانس جامعه است. اما در صورت استاندارد شدن $X_۱$ و $X_۲$ سهم یکسانی به $Z_۱$ و $Z_۲$ در ساختار مولفه‌های اصلی $\boldsymbol{\rho}$ داده می‌شود. به کمک رابطه (۷-۳-۱)

$$\rho_{Y_۱,Z_۱} = e_{۱۱}\sqrt{\lambda_۱} = ۰.۷۰۷\sqrt{۱.٤} = ۰.۸۳۷$$

$$\rho_{Y_۱,Z_۲} = e_{۲۱}\sqrt{\lambda_۲} = ۰.۷۰۷\sqrt{۱.٤} = ۰.۸۳۷$$

واریانس نسبی مولفه اول $\frac{\lambda_۱}{p} = \frac{۱.٤}{۲} = ۰.۷$ از کل واریانس جامعه است.

مثال (۶-۳-۱) نشان می‌دهد که مولفه‌های اصلی حاصله از Σ با مولفه های اصلی از $\boldsymbol{\rho}$ کاملا متفاوت است. همچنین مولفه‌های اصلی تابعی ساده از یکدیگر نیستند و این تاثیر استاندارد کردن متغیرها را نشان می دهد. معمولا وقتی واحد اندازه گیری متغیرها متفاوت باشد از متغیرهای استاندارد شده استفاده می‌شود که اینکار با به کاربردن ماتریس ضرائب همبستگی انجام می‌شود.

٤-۷ مولفه‌های اصلی براساس ماتریس کوواریانس نمونه

معمولا در عمل ماتریس کوواریانس Σ مجهول است. بنابراین مؤلفه های اصلی جمعیت بی فایده خواهند بود و تصمیم گیری در این مورد که کدام یک از مؤلفه های اصلی دارای واریانس های به اندازه کافی کوچک هستند که بتوان از آن صرف نظر کرد، باید بر مبنای نمونه صورت گیرد.

فرض کنید $\boldsymbol{X}_۱, \ldots, \boldsymbol{X}_n$ یک نمونه تصادفی از جامعه نرمال p بعدی با میانگین $\boldsymbol{\mu}$ و ماتریس کوواریانس Σ باشد. با برآورد برآورد درستنمایی ماکسیمم $\boldsymbol{\mu}$ و Σ به صورت زیر:

$$\hat{\boldsymbol{\mu}} = \overline{\boldsymbol{X}}$$

$$\hat{\Sigma} = \frac{۱}{n}\sum_{i=۱}^{n}(\boldsymbol{X}_i-\overline{\boldsymbol{X}})\,(\boldsymbol{X}_i-\overline{\boldsymbol{X}})' = \boldsymbol{S}_n$$

کار را دنبال می‌کنیم. مقادیر ویژه و بردارهای ویژه $\boldsymbol{S}_n$ برآوردر درستنمایی برای همتای آنها درΣ می‌باشند. به عبارت دیگر زوج$(\hat{\lambda}_i, \hat{\boldsymbol{e}}_i)$ مقدار و بردار ویژه از $\boldsymbol{S}_n$ برآورد درستنمایی ماکسیم $(\hat{\lambda}_i, \hat{\boldsymbol{e}}_i)$ در Σ هستند. پس مولفه های اصلی iام نمونه

$$\hat{Y}_i = \hat{\boldsymbol{e}}_i' \boldsymbol{X} = \hat{e}_{1i}X_1 + \hat{e}_{2i}X_2 + \cdots + \hat{e}_{pi}X_p \qquad i = 1,2,\dots,p$$

که در آن $\boldsymbol{X}$ هر مشاهده روی متغیرهای تصادفی $\boldsymbol{X}_1, \dots, \boldsymbol{X}_n$ است. همچنین $\hat{\lambda}_j$ برآورد واریانس نمونه ای $\hat{Y}_j$ و برای هر $cov(\hat{Y}_i, \hat{Y}_j) = 0,\ i \neq j$ است. بعلاوه

$$\text{کل واریانس نمونه} = \sum_{i=1}^{p} S_{ii} = \hat{\lambda}_1 + \cdots + \hat{\lambda}_p$$

$$\boldsymbol{r}_{\hat{Y}_i, X_k} = \hat{\boldsymbol{e}}_{ki}\sqrt{\frac{\hat{\lambda}_i}{S_{kk}}} \qquad i,k = 1,2,\dots,p$$

همچنین مشابه بخش قبل می‌توان مولفه‌های اصلی برحسب متغیرهاس استاندارد شده و R محاسبه کرد.

مثال ۷-٤-۱: از بررسی یک نمونه‌ی١٤ تایی بر روی ٥ متغیر اقتصادی اجتماعی نتایج حاصل شده است

$$\overline{\boldsymbol{X}}' = (4.32 \quad 14.01 \quad 1.95 \quad 2.17 \quad 2.45)$$

$$\boldsymbol{S}_{14} = \begin{bmatrix} 4.308 & 1.683 & 1.803 & 2.155 & -0.253 \\ & 1.768 & 0.588 & 0.177 & 0.176 \\ & & 0.801 & 1.065 & -0.158 \\ & & & 1.970 & -0.357 \\ & & & & 0.504 \end{bmatrix}$$

آیا می‌توان تغییرات را در یک یا دو مولفه اصلی خلاصه نمود؟ نتایج در جدول زیر خلاصه شده است.

متغیر	$\hat{\boldsymbol{e}}_1(r_{\hat{Y}_1,X_j})$	$\hat{\boldsymbol{e}}_2\left(r_{\hat{Y}_2,X_j}\right)$	$\hat{\boldsymbol{e}}_3$	$\hat{\boldsymbol{e}}_4$	$\hat{\boldsymbol{e}}_5$
X_1	$0.781(0.99)$	$-0.071(-0.04)$	0.004	0.542	-0.302
X_2	$0.306(0.61)$	$-0.764(-0.76)$	-0.162	-0.545	-0.010
X_3	$0.334(0.98)$	$0.083(-0.12)$	0.015	0.05	0.937
X_4	$0.426(0.80)$	$-0.579(0.55)$	0.220	-0.636	-0.173
X_5	$-0.054(-0.20)$	$-0.262(-0.49)$	0.962	-0.051	0.024
λ_i	6.639	1.786	0.390	0.230	0.014
درصد تجمعی واریانس کل	٧٤.١	٩٣.٢	٩٧.٤	٩٩.٩	١٠٠

بنابر اطلاعات فوق دو مولفه اصلی کافی است، یعنی ۱٤ مشاهده ٥ متغیر به مشاهده دو متغیر مبدل می‌گردد.

تمرین ٥-۷

۱- فرض کنید Σ ماتریس کوواریانس بردار تصادفی $\boldsymbol{X}' = (X_1 \quad X_2 \quad \dots \quad X_p)$ بوده و ماتریس کوواریانس Σ دارای مقادیر ویژه بردارهای ویژه $(\lambda_i, \boldsymbol{e}_i) \;\; i = 1,2,\dots,p$ باشد و $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$، مولفه اصلی iام را بوسیله‌ی

$$Y_i = \mathbf{e}_i' \mathbf{X} = e_{i1}X_1 + \dots + e_{pi}X_p \qquad\qquad i = 1,2,\dots,p$$

مشخص می کنیم. ثابت کنید با این انتخاب‌ها، داریم

$$Var(Y_i) = \mathbf{e}_i'\Sigma\, \mathbf{e}_i = \lambda_i \qquad\qquad i = 1,2,\dots,p$$

$$Cov(Y_i, Y_k) = \boldsymbol{e}_i'\Sigma\, \boldsymbol{e}_k = 0 \qquad\qquad i \neq k$$

۲- نشان دهید مجذور همبستگی چندگانه بین X_i و $Y_1, \dots, Y_k$ به صورت زیر محاسبه می‌شود

$$\rho^2_{x_{i(y_1,\dots,y_k)}} = \sum_{m=1}^{k} \lambda_m\, e^2_{im}/\sigma_{ii}$$

۳- در تمرین ۲ همبستگی چندگانه برای مثال (۷-۲-۱) را بدست آورید و تفسیر کنید.

٤- برای $\Sigma = diag\left(\sigma_{11}, \dots, \sigma_{pp}\right)$ مولفه های اصلی بدست آورید و تفسیر کنید.

٥- برای $\Sigma = \sigma^2[(1-\rho)\boldsymbol{I_p} + \rho\boldsymbol{LL}']$ نشان دهید:

الف- بزرگترین مقدار ویژه Σ عبارت است از $\lambda_1 = \sigma^2(1 + (p-1)\rho)$

ب- اولین مولفه اصلی $\mathbf{X}$ عبارت است از $Z_1 = \frac{1}{\sqrt{p}}\sum_{i=1}^{p} X_i$

٦- فرض کنید $\boldsymbol{X}' = (X_1 \quad X_2 \quad X_3 \quad X_4)$ یک بردار تصادفی با ماتریس کوواریانس زیر باشد:

$$\Sigma = \begin{pmatrix} a & b & d & c \\ & a & c & d \\ & & a & b \\ & & & a \end{pmatrix}$$

نشان دهید که مولفه های اصلی $\mathbf{X}$ عبارتند از:

$$Z_1 = \frac{1}{2}(X_1 + X_2 + X_3 + X_4)$$

$$Z_2 = \frac{1}{2}(X_1 + X_2 - X_3 - X_4)$$

$$Z_3 = \frac{1}{2}(X_1 - X_2 + X_3 - X_4)$$

$$Z_4 = \frac{1}{2}(X_1 - X_2 - X_3 + X_4)$$

۶-۷ حل تمرین

۱- برای $i = 1,2,\dots,p$ داریم:

$$Var(Y_i) = Var(\boldsymbol{e}_i'\boldsymbol{X}) = \boldsymbol{e}_i' Cov(\boldsymbol{X})\boldsymbol{e}_i = \boldsymbol{e}_i'\Sigma\,\boldsymbol{e}_i = \boldsymbol{e}_i'\lambda_i\,\boldsymbol{e}_i = \lambda_i$$

و برای هر $i \neq k$

$$\begin{aligned} Cov(Y_i, Y_k) &= Cov(\boldsymbol{e}_i'\boldsymbol{X}, \boldsymbol{e}_k'\boldsymbol{X}) \\ &= \boldsymbol{e}_i'\,Cov(\boldsymbol{X})\boldsymbol{e}_k \\ &= \boldsymbol{e}_i'\ \Sigma\ \boldsymbol{e}_k \\ &= \boldsymbol{e}_i'\ \lambda_k\ \boldsymbol{e}_k \\ &= 0 \end{aligned}$$

۲- بردار تصادفی $(X_i \quad Y_1 \quad \dots \quad Y_k)$ را در نظر بگیرید. ماتریس کوواریانس این بردار تصادفی عبارت است از:

$$\boldsymbol{M} = \begin{pmatrix} \sigma_{ii} & & A & & \\ & \lambda_1 & 0 & \dots & 0 \\ & 0 & \lambda_2 & \dots & \vdots \\ B & \vdots & \vdots & \ddots & \vdots \\ & 0 & 0 & \dots & \lambda_k \end{pmatrix}$$

که در آن A و B ترانهاده یکدیگر هستند و برابرند با :

$$B = Cov(X_i, \begin{pmatrix} \boldsymbol{e}_1'\boldsymbol{X} \\ \vdots \\ \boldsymbol{e}_k'\boldsymbol{X} \end{pmatrix})$$

$$= \begin{pmatrix} Cov(X_i, \boldsymbol{e}_1'\boldsymbol{X}) \\ \vdots \\ Cov(X_i, \boldsymbol{e}_k'\boldsymbol{X}) \end{pmatrix} = \begin{pmatrix} \boldsymbol{e}_1'\boldsymbol{\sigma}_i \\ \vdots \\ \boldsymbol{e}_k'\boldsymbol{\sigma}_i \end{pmatrix} = \begin{pmatrix} \lambda_1 e_{i1} \\ \vdots \\ \lambda_k e_{ik} \end{pmatrix} = \boldsymbol{\sigma}_{21}$$

که در آن $\boldsymbol{\sigma}_j$، به معنی ستون j ام از ماتریس Σ است. حال می‌توانیم ماتریس کوواریانس $(X_i \quad Y_1 \quad \dots \quad Y_k)$ دوباره بازنویسی کنیم:

$$\boldsymbol{M} = \begin{pmatrix} \sigma_{ii} & \lambda_1 e_{i1} & & \dots & \lambda_k e_{ik} \\ \lambda_1 e_{i1} & \lambda_1 & 0 & \dots & 0 \\ \vdots & 0 & \lambda_2 & \dots & \vdots \\ \lambda_k e_{ik} & \vdots & \vdots & \ddots & \vdots \\ & 0 & 0 & \dots & \lambda_k \end{pmatrix}$$

با تناظر قراردادن کمیت ها با فرمول های مرتبط در افراز ماتریس کوواریانس $\Sigma = \begin{pmatrix} \sigma_{11} & \boldsymbol{\sigma}_{21}' \\ \boldsymbol{\sigma}_{21} & \Sigma_{22} \end{pmatrix}$، ضریب همبستگی چندگانه را به دست می آوریم.

$$\boldsymbol{\sigma}'_{۲۱}\Sigma_{۲۲}^{-۱}\boldsymbol{\sigma}_{۲۱} = (\lambda_۱ e_{i۱} \dots \lambda_k e_{ik}) diag(\frac{۱}{\lambda_۱}, \dots, \frac{۱}{\lambda_k}) \begin{pmatrix} \lambda_۱ e_{i۱} \\ \vdots \\ \lambda_k e_{ik} \end{pmatrix}$$

$$= \sum_{m=۱}^{k} \lambda_m e_{im}^{۲}$$

نتیجه می‌شود:

$$\rho^{۲}_{x_{i(y_۱,\dots,y_k)}} = \sum_{m=۱}^{k} \lambda_m e_{im}^{۲}/\sigma_{ii}$$

۳- بنابر اطلاعات مثال ۷-۲-۱ مقادیر ویژه و بردارهای ویژه ماتریس کوواریانس عبارت است از($\sigma_{۱۱} = ۱$، $\sigma_{۲۲} = ۵$، $\sigma_{۳۳} = ۲$)

$e_{i۱}$	-۰.۳۸۳	۰.۹۲٤	۰	$\lambda_۱ = ۵.۸۳$
$e_{i۱}^{۲}$	۰.۱٤۷	۰.۸۵۳	۰	
$e_{i۲}$	۰	۰	۱	$\lambda_۲ = ۲$
$e_{i۲}^{۲}$	۰	۰	۱	
$e_{i۳}$	۰.۹۲٤	۰.۳۸۳	۰	$\lambda_۳ = ۰.۱۷$
$e_{i۳}^{۲}$	۰.۸۵۳	۰.۱٤۷	۰	

با قرار دادن در فرمول $\rho^{۲}_{x_{i(y_۱,\dots,y_k)}} = \sum_{m=۱}^{k} \lambda_m e_{im}^{۲}/\sigma_{ii}$ مقادیرهمبستگی چندگانه عبارت است از:

$\rho_{X_۱(Y_۱)} = ۰.۸۵۷$ $\quad$ $\rho_{X_۱(Y_۱,Y_۲)} = ۰.۸۵۷$ $\quad$ $\rho_{X_۱(Y_۱,Y_۲,Y_۳)} = ۱$

$\rho_{X_۲(Y_۱)} = ۰.۹۹٤$ $\quad$ $\rho_{X_۲(Y_۱,Y_۲)} = ۰۹۹٤$ $\quad$ $\rho_{X_۲(Y_۱,Y_۲,Y_۳)} = ۱$

$\rho_{X_۳(Y_۱)} = ۰$ $\quad$ $\rho_{X_۳(Y_۱,Y_۲)} = ۱$ $\quad$ $\rho_{X_۳(Y_۱,Y_۲,Y_۳)} = ۱$

٤- اگر مقادیر روی قطر اصلی ماتریس کوواریانس از بزرگ به کوچک مرتب کنیم متغیرهای متناظر، مولفه های اصلی را تشکیل می‌دهد.

۵- قرار دادن $\sigma^۲ = ۱$. از حل معادله مشخصه ی زیر، مقادیر ویژه به دست می آید.

$$\left|\Sigma - \lambda \boldsymbol{I}_p\right| = ۰$$

یا

$$\left|(۱-\rho)\boldsymbol{I}_p + \rho \boldsymbol{LL}' - \lambda \boldsymbol{I}_p\right| = ۰$$

با فاکتور گیری داریم:

$$\left|p\rho \frac{۱}{p}\boldsymbol{LL}' - \big(\lambda - (۱-\rho)\big)\boldsymbol{I}_p\right| = ۰$$

$\frac{۱}{p}\boldsymbol{LL}'$ یک ماتریس خودتوان است پس دارای مقادیر ویژه صفر ویک است. بنابر این یک مرتبه مقدار ویژه برابر است با

$\lambda-(۱-\rho)=p\rho$ و $\lambda-(۱-\rho)$ $(p-۱)$ مرتبه مقدار ویژه برابر است با صفر یعنی $\lambda-(۱-\rho)=۰$ به عبارت دیگر مقادیر ویژه عبارتند از:

$$\begin{cases}\lambda_۱=۱-\rho+p\rho=۱+(p-۱)\rho\\ \lambda_۲=۱-\rho\end{cases}$$

واضح است که بزرگی هر یک به مقدار ضریب همبستگی وابسته است. در صورتی که $۰<\rho$ باشد آنگاه $\lambda_۱<\lambda_۲$.

ب- برای تعیین بردار ویژه معادله زیر باید حل شود:

$$\Sigma\boldsymbol{e}=\lambda_۱\boldsymbol{e}$$

$$\left((۱-\rho)\boldsymbol{I}_p+\rho\boldsymbol{L}\boldsymbol{L}'-\lambda\boldsymbol{I}_p\right)\boldsymbol{e}=\lambda_۱\boldsymbol{e}$$

$$(۱-\rho)\boldsymbol{e}+\rho\boldsymbol{L}\boldsymbol{L}'\boldsymbol{e}=\lambda_۱\boldsymbol{e}$$

اگر $\boldsymbol{e}'=(e_۱\ \dots\ e_p)$ باشد، از طرفی $۱-\rho-\lambda_۱=p\rho$،

$$-p\rho\boldsymbol{e}+\rho\boldsymbol{L}\sum_{j=۱}^{p}e_j=۰$$

بنابراین $i=۱,۲,\dots,p$ $\quad pe_i=-\sum_{j=۱}^{p}e_j$ در نتیجه $e_۱=\cdots=e_p$ پس

$$\boldsymbol{e}=\frac{۱}{\sqrt{p}}\begin{pmatrix}۱\\ \vdots\\ ۱\end{pmatrix}=\frac{۱}{\sqrt{p}}\boldsymbol{L}$$

و اولین مولفه اصلی به صورت زیر مبدل می گردد:

$$Y_۱=\frac{۱}{\sqrt{p}}\boldsymbol{L}'\boldsymbol{X}=\frac{۱}{\sqrt{p}}\sum_{i=۱}^{p}X_i$$

٦- دترمینال ماتریس Σ عبارت است از:

$$|\Sigma|=(a-b-c+d)(a+b+c+d)(a+b-c-d)(a-b+c-d)$$

است پس مقادیر ویژه Σ عبارتند از:

$$\lambda_۱=a+b+c+d$$

$$\lambda_۲=a+b-c-d$$

$$\lambda_۳=a-b+c-d$$

$$\lambda_٤=a-b-c+d$$

در مرحله بعدی نشان دهید برای

$$P=\frac{۱}{۲}\begin{pmatrix}۱&۱&۱&۱\\ ۱&۱&-۱&-۱\\ ۱&-۱&-۱&۱\\ ۱&-۱&۱&-۱\end{pmatrix}$$

و $D=diag(\lambda_۱,\lambda_۲,\lambda_۳,\lambda_٤)$ داریم: $\Sigma=PDP'$.

۷-۷ سوالات تستی

۱- در مباحث مولفه های اصلی مقدار $tr\Sigma$ با کدام کمیت برابر نیست ؟ $\boldsymbol{e}' = (۱,۱,\dots,۱)$ و $D_\lambda = diag(\lambda_۱,\dots,\lambda_p)$

الف. $\sum_{i=۱}^{p}\sigma_{ii}$ ب. $\sum_{i=۱}^{p}\lambda_i$ ج. $\boldsymbol{e}'\Sigma\boldsymbol{e}$ د. $\boldsymbol{e}'D_\lambda\boldsymbol{e}$

۲- اگر همه مقادیر ویژه Σ با هم برابر λ باشند مقدار همبستگی بین X_i و Y_j چقدر است؟$(i \neq j)$

الف. ۱ ب. $y_{ij}\sqrt{\frac{\lambda}{\sigma_{ii}}}$ ج. $y_{ij}\sqrt{\frac{\sigma_{ii}}{\lambda}}$ د. صفر

۳- فرض $\boldsymbol{X}\sim N_۲\left(\boldsymbol{\mu}\ ,\ \begin{bmatrix} ٤ & ۱ \\ ۱ & ٤ \end{bmatrix}\right)$ کنید مولفه های اصلی کدام هستند؟

الف. $\begin{cases} Y_۱ = \frac{۱}{\sqrt{۲}}X_۱ - \frac{۱}{\sqrt{۲}}X_۲ \\ Y_۲ = \frac{۱}{\sqrt{۲}}X_۱ + \frac{۱}{\sqrt{۲}}X_۲ \end{cases}$ ب. $\begin{cases} Y_۱ = X_۲ \\ Y_۲ = \frac{۱}{\sqrt{۲}}X_۱ + \frac{۱}{\sqrt{۲}}X_۲ \end{cases}$

ج. $\begin{cases} Y_۱ = X_۱ + X_۲ \\ Y_۲ = X_۱ - X_۲ \end{cases}$ د. $\begin{cases} Y_۱ = \frac{۱}{\sqrt{۲}}X_۱ - \frac{۱}{\sqrt{۲}}X_۲ \\ Y_۲ = X_۱ - X_۲ \end{cases}$

٤- کدامیک از گزینه های زیر در مورد مولفه های اصلی نادرست است؟

الف. تحت تبدیلات خطی پایدار نمی باشند.

ب. از یکدیگر ناهمبسته هستند.

ج. تحت تبدیلات خطی پایدار هستند.

د. از یکدیگر به طور مجانبی مستقل هستند.

۵- اگر ماتریس کوواریانس $\Sigma = \begin{pmatrix} ۱ & ۰.۲۵ & ۰.۲۵ \\ & ۱ & ۰.۲۵ \\ & & ۱ \end{pmatrix}$ واریانس دومین مولفه اصلی چقدر است؟

الف. ۰.۲۵ ب. ۲ ج. ۱.۵ د. ۰.۷۵

۷- ۸ پاسخ سوالات تستی

۱- ج ۲- د ۳- الف ٤- الف ۵- د