همبستگی متعارف استوار

محمد جعفری - مسعود یارمحمدی

دانشگاه پیام نور، گروه علمی آمار تهران، ایران

چکیده: تحلیل همبستگی متعارف یک روش آماری چندمتغیره به منظور بررسی همبستگی بین یک ترکیب خطی از متغیرهای دو مجموعه مجزا بر اساس ماتریس واریانس کوواریانس (یا همبستگی) میان متغیرها معرفی شده است. حال از آنجایی که این ماتریس‌ها به شدت به داده‌های دورافتاده حساس می‌باشند، نتایج به دست آمده از آن‌ها در حضور داده‌های دورافتاده قابل اعتماد نمی‌باشند. به منظور استوارسازی تحلیل همبستگی متعارف، سه شاخص مختلف بر اساس روش تعقیب تصویر و همچنین روش رگرسیون متناوب استوار معرفی شده است.

واژه‌های کلیدی: همبستگی متعارف استوار، تعقیب تصویر، کوواریانس استوار، رگرسیون متناوب استوار

۱ تحلیل همبستگی متعارف

هدف تحلیل همبستگی متعارف، شناسایی و کمی نمودن رابطه‌ی بین دو مجموعه از متغیرها می‌باشد. این روش، همبستگی بین یک ترکیب خطی مانند $U = a'X^{(1)}$ از متغیرهای مجموعه $X^{(1)}_{p\times 1}$ با یک ترکیب خطی مانند $V = b'X^{(2)}$ از متغیرهای مجموعه $X^{(2)}_{q\times 1}$ $(p \leq q)$ را مورد بررسی قرار می‌دهد. ابتدا زوج ترکیب خطی که بیشترین همبستگی دارند را معین نموده که $\max\limits_{a,b} \mathrm{Corr}(U_1, V_1) = \rho^*_1$ می‌باشد، سپس زوج ترکیب خطی بعدی که بزرگترین همبستگی را بین همه‌ی زوج‌ها داشته به طوری که با زوج انتخاب شده قبلی ناهمبسته باشد، انتخاب شده و این فرایند به همین صورت ادامه می‌یابد و $\mathrm{Corr}(U_k, V_k) = \rho^*_k$ به ازای $k = 2, 3, \ldots, p$ به دست می‌آید که به $\rho^{*2}_1 \geq \rho^{*2}_2 \geq \ldots \geq \rho^{*2}_p$ منجر می‌شود (جانسون و ویچرن[1]، ۲۰۰۰).

[1] Johnson & Wichern

۲ استوارسازی

از آنجایی که ماتریس کوواریانس (همبستگی) و در نتیجه مقادیر و بردارهای ویژه متناظر با آن‌ها به شدت به داده‌های دورافتاده حساس می‌باشند، لذا نتایج به دست آمده در حضور داده های دورافتاده، استوار نخواهند بود. برای استوار نمودن روش‌های کلاسیک، کرنل[۲] (۱۹۹۱) M– برآوردگرها و سپس کراکس و دهون [۳] (۲۰۰۲) برآوردگرهای MCD را پیشنهادنمودند. روش‌های استوار قادرند به مشاهدات وزن‌های نابرابر اختصاص دهند. برای انجام تحلیل همبستگی متعارف استوار از روش تعقیب تصویر [۴] (PP) استفاده می‌شود، که با جایگزینی مقدار ضریب همبستگی در رابطه‌ی $\mathrm{Corr}(U_k, V_k) = \rho_k^*$ به وسیله یک اندازه از ضریب همبستگی استوار، همبستگی متعارف را استوار می‌نماید. این روش، در جستجوی دو مجموعه a و b می‌باشد که همبستگی بین متغیرهای x و y تصویر شده روی این مجموعه‌ها را ماکسیمم کند. روش دیگر برای تحلیل همبستگی متعارف استوار، استفاده از روش رگرسیون‌متناوب‌استوار می‌باشد که توسط والد [۵] (۱۹۶۶) پیشنهاد شده است.

۱.۲ روش تعقیب تصویر

سه شاخص تصویر (PI) در این مقاله ارائه شده است که عبارتند از:

۱.۱.۲ شاخص تصویری M– برآوردگر دو متغیره $(PP-M)$

در همبستگی متعارف، متغیرتصادفی دوبعدی $z=(u,v)'$ مفروض‌است. M– برآوردگر دو متغیره [۶] مارونا (۱۹۷۶) با پارامتر مکان $\mu(z)$ و ماتریس پراکنش $C(z)$ از حل معادلات زیر حاصل می‌شود:

$$\mu = \frac{E\left[w_{1}\left((z-\mu)'C^{-1}(z-\mu)\right)z\right]}{E\left[w_{1}\left((z-\mu)'C^{-1}(z-\mu)\right)\right]}$$
$$C = E\left[w_{2}\left((z-\mu)'C^{-1}(z-\mu)\right)(z-\mu)(z-\mu)'\right]$$

که μ یک بردار دو متغیره، C یک ماتریس مثبت متقارن و w_1 و w_2 توابع وزنی مشخصی هستند (برانکو و همکاران، ۲۰۰۵).

---
۲ Karnel
۳ Croux & Dehon
۴ Projection Pursuit
۵ Wold
۶ Brivariate M estimator

۲.۱.۲ شاخص تصویری برآوردگر MCD دو متغیره $(PP-MCD)$

به جای $M-$ برآوردگر دو متغیره که تحت مقادیر بزرگ آلوده، استواری خود را از دست می‌دهد، برآوردگر ماتریس کوواریانس با کمترین دترمینان (MCD) انتخاب می‌شود (روسیو ۱۹۸۵، و روسیو و ون درسن [۷] ۱۹۹۹). در حالت معمولی برای استوار سازی تحلیل همبستگی متعارف، Σ به صورت استوار برآورد شده، مقادیر و بردارهای ویژه از نسخه برآورد شده $\Sigma_{11}^{-\frac{1}{2}}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-\frac{1}{2}}$ یا $\Sigma_{22}^{-\frac{1}{2}}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-\frac{1}{2}}$ محاسبه می‌شوند. می‌توانیم $M-$ برآوردگر را به عنوان برآوردگر کوواریانس استوار مانند کرنل (۱۹۹۱) استفاده کنیم، اما همانگونه که برانکو و همکاران (۲۰۰۵) نشان داده‌اند، این برآوردگرها دارای خاصیت استواری کمی در ابعاد بالاتر هستند که این موضوع در بخش ۳ مورد بررسی قرار می‌گیرد. برآوردگر MCD با جستجو در زیرمجموعه‌ای با اندازه h از داده‌هایی به دست می‌آید که دارای ماتریس کوواریانس تجربی محاسبه شده با کوچکترین دترمینان می‌باشد که در این حالت، بیشترین استواری در $h \approx \frac{n}{2}$ یا $h \approx \frac{3n}{4}$ به دست می‌آید.

۳.۱.۲ شاخص تصویری ضریب همبستگی رتبه‌ای اسپیرمن $(PP-SPM)$

اندازه ارتباط متقابل معروف $\rho_s(u,v)$، به عنوان همبستگی بین رتبه‌های دومتغیر تصادفی u و v تعریف می‌شود. این معیار، خاصیت ناپارامتری دارد و به شرایط تقارن توزیع نرمال تکیه ندارد. وقتی از یک توزیع نرمال دو متغیره نمونه می‌گیریم، ρ_s کمیتی مشابه ضریب همبستگی پیرسون برآورد نمی‌کند (برانکو وهمکاران، ۲۰۰۵). برای مقایسه برآورد همبستگی متعارف با استفاده از ضریب همبستگی رتبه‌ای اسپیرمن با روش‌های دیگر، تبدیل $2\sin\left(\frac{\pi\rho_s}{6}\right)$ را انجام می‌دهیم تا برآورد تحت نرمال بودن، سازگار شود (برانکو وهمکاران، ۲۰۰۵).

شاخص‌های تصویری معرفی شده در بخش‌های ۱.۱.۲ تا ۳.۱.۲ برای یافتن اولین زوج بردارهای متعارف از رابطه $\max\limits_{a,b}\operatorname{Corr}(U_1,V_1)=\rho_1^*$ استفاده می‌نمایند. با انجام دادن شرایط

$$\operatorname{cov}(u_L,u_j)=\alpha_L'\Sigma_{xx}\alpha_j=\operatorname{cov}(v_L,v_j)=\beta_L'\Sigma_{yy}\beta_j=\delta_{Lj}$$

$$\delta_{Lj}=\begin{cases}1 & L=j\\ 0 & 1\leq j<L\end{cases}$$

[۷] Rousseeuw & Van Driessen

برای محاسبه بردارهای متعارف مراتب بالاتر پیشنهاد می‌شود، کار را با برآورد کردن Σ با استفاده از برآوردگرهای استوارتر، مثل برآوردگر MCD دووزنی، شروع کنیم (روسیو و ون درسن، ۱۹۹۹). سپس Σ افراز می‌شود و با یک تجزیه طیفی Σ_{11} و Σ_{22} به $\Sigma_{11} = UKU'$ و $\Sigma_{22} = VLV'$ منجر می‌شوند که K و L ماتریس قطری و U و V ماتریس‌های متعامد هستند. اکنون متغیرهای اصلی تبدیل یافته عبارتند از:

$$(x^*, y^*) = \left(K^{-\frac{1}{2}}U'x, L^{-\frac{1}{2}}V'y\right)$$

متغیرهای جدید (x^*, y^*) همبستگی متعارف مشابه متغیرهای اصلی (x, y) دارند و ضرایب متعارف برای $L = 1, \ldots, p$ در روابط زیر صدق می‌کنند:

$$a_L = UK^{-\frac{1}{2}}a_L^* \quad , \quad b_L = VL^{-\frac{1}{2}}b_L^*$$

برای یافتن اولین بردارهای متعارف a_1^* و b_1^* شاخص تصویر $PI(a'x^*, b'y^*)$ باید تحت شرایط $\text{Var}(a'x^*) = a'a = 1$ و $\text{Var}(b'y^*) = b'b = 1$ ماکسیمم شود. با استفاده از روابط بازگشتی در مختصات قطبی (برانکو و همکاران، ۲۰۰۵) a_1^* و b_1^* به دست می‌آیند. اکنون با استفاده از $(L-1)$ زوج ضرایب متعارف (a_j^*, b_j^*) به ازای $j = 1, \ldots, L-1$ می‌خواهیم زوج L اُم $(2 \leq L \leq p)$ را برآورد کنیم.

توجه داشته باشید که شرایط $a_L'\Sigma_{11}a_j = 0$ و $b_L'\Sigma_{22}b_j = 0$ برای $j \neq L$ به شرایط متعامد $a_L'^*a_j^* = 0$ و $b_L'^*b_j^* = 0$ تبدیل می‌شوند. ابتدا دو ماتریس متعامد $A = [a_1^* \ldots a_{L-1}^* \mid A^r]$ و $B = [b_1^* \ldots b_{L-1}^* \mid B^r]$ را ایجاد می‌کنیم که به ترتیب A^r و B^r یکامتعامد بر اساس زیرفضای متعامد روی $a_1^*, \ldots, a_{L-1}^*$ و $b_1^*, \ldots, b_{L-1}^*$ می‌باشند. سپس متغیرها را روی زیرفضای متعامد، برای بردارهای متعارف تصویر می‌کنیم:

$$(x^{**}, y^{**}) = ((A^r)'x^*, (B^r)'y^*)$$

اکنون باید a^{**} و b^{**} را بیابیم که تحت شرط این که نُرم a^{**} و b^{**} برابر یک باشند، $PI(a'x^{**}, b'y^{**})$ را ماکسیمم نماید. همانند قبل، $a_L^* = A^r a^{**}$ و $b_L^* = B^r b^{**}$ به دست می‌آیند. بعد از به دست آوردن بردارهای متعارف (x^*, y^*)، با استفاده از تبدیل زیر، متغیرهای اصلی محاسبه می‌شوند:

$$(a_L, b_L) = \left(UK^{-\frac{1}{2}}a_L^*, VL^{-\frac{1}{2}}b_L^*\right) \quad , \quad L = 2, \ldots, p$$

در نهایت، همبستگی‌های متعارف با محاسبه $\rho_L = PI(u_L, v_L)$ برآورد می‌شوند که $(u_L, v_L) = (a_L'x, b_L'y)$ برای $L = 1, \ldots, p$ می‌باشد.

۲.۲ روش رگرسیون متناوب

در این روش، فرض کنید یک مقدار آغازین برای b داشته‌باشیم که‌برابر β قرار می‌دهیم. آن‌گاه مسئله ماکسیمم‌سازی رابطه $\mathrm{Corr}(U_k, V_k) = \rho_k^*$ به صورت زیر ساده می‌شود:

$$\alpha = \underset{a}{\mathrm{argmax}}\mathrm{Corr}(a'x, \beta'y)$$

که از رگرسیون چندگانه پیروی می‌کند. α، متناسب با ضرایب رگرسیونی a در مدل زیر می‌باشد:

$$\beta'y = a'x + \gamma_۱ + \varepsilon_۱$$

به طور مشابه، برای یک α ی ثابت، β متناسب با ضرایب رگرسیونی b در مدل زیر می‌باشد:

$$\alpha'x = b'y + \gamma_۲ + \varepsilon_۲$$

این موضوع به طرح رگرسیون متناوبی منجر می‌شود که برای انجام آن، به ترتیب گام‌های ذیل را طی می‌نماییم.

گام۱: در نظر گرفتن مقدار آغازین $X_\circ = X - ۱\bar{x}'$ و $Y_\circ = Y - ۱\bar{y}'$

گام۲: برای $L = ۱, \ldots, k$:

گام ۲-۱: فضای مانده‌ها (فقط اگر $L > ۱$):

$$X_{L-۱} = \left(I_n - \frac{u_{L-۱}u'_{L-۱}}{u'_{L-۱}u_{L-۱}}\right) X_{L-۲} \quad , \quad Y_{L-۱} = \left(I_n - \frac{v_{L-۱}v'_{L-۱}}{v'_{L-۱}v_{L-۱}}\right) Y_{L-۲}$$

گام ۲-۲: مقادیر آغازین (با استفاده از اولین مؤلفه اصلی $z_۱^{L-۱}$ از $X_{L-۱}$):

$$\hat{b}_L^{(\circ)} = \left(Y'_{L-۱}Y_{L-۱}\right)^{-۱} Y'_{L-۱}z_۱^{L-۱} \Rightarrow \beta_L^{(\circ)} = \frac{\hat{b}_L^{(\circ)}}{\|\hat{b}_L^{(\circ)}\|} \Rightarrow v_L^{(\circ)} = Y_{L-۱}\beta_L^{(\circ)}$$

گام ۲-۳: از $s = ۱$ تا همگرایی:

$$\hat{a}_L^{(s)} = \left(X'_{L-۱}X_{L-۱}\right)^{-۱} X'_{L-۱}v_L^{s-۱} \Rightarrow \alpha_L^{(s)} = \frac{\hat{a}_L^{(s)}}{\|\hat{a}_L^{(s)}\|} \Rightarrow u_L^{(s)} = X_{L-۱}\alpha_L^{(s)}$$

$$\hat{b}_L^{(s)} = \left(Y'_{L-۱}Y_{L-۱}\right)^{-۱} Y'_{L-۱}u_L^{s-۱} \Rightarrow \beta_L^{(s)} = \frac{\hat{b}_L^{(s)}}{\|\hat{b}_L^{(s)}\|} \Rightarrow v_L^{(s)} = Y_{L-۱}\beta_L^{(s)}$$

گام ۲-۴: بعد از همگرایی، نتایج در u_L^*، v_L^*، $\alpha_۱^*$، $\beta_۱^*$:

$$|r_L = \mathrm{Corr}(u_L^*, v_L^*)|$$

گام ۲-۴-۱: اگر $L = 1$:

$$u_1 = u_1^* \ , \ v_1 = v_1^* \ , \ \hat{\alpha}_1 = \alpha_1^* \ , \ \hat{\beta}_1 = \beta_1^*$$

گام ۲-۴-۲: اگر $L > 1$:

$$u_L = u_L^* \quad \Rightarrow \quad \hat{\alpha}_L = (X_0'X_0)^{-1}X_0'u_L \quad , \quad v_L = v_L^* \quad \Rightarrow \quad \hat{\beta}_L = (Y_0'Y_0)^{-1}Y_0'v_L$$

۱.۲.۲ رگرسیون متناوب استوار

روش رگرسیون متناوب استوار[۸] (RAR) شامل همه برآوردها با استفاده از رگرسیون استوار به جای رگرسیون حداقل مربعات می‌باشد. یک برآوردگر مفید، برآوردگر حداقل مربعات پیراسته [۹] (LTS) (روسیو، ۱۹۸۴) می‌باشد که دارای استواری بالایی بوده و محاسبه ن آسان می‌باشد.

از آن‌جایی که رگرسیون‌ها متناوب می‌باشند، برای RAR باید مشاهدات دورافتاده در فضای x و فضای y وزن‌دار شوند. برای استخراج متغیرهای متعارف L اُم ($L = 1,\ldots,k$)، باید وزن‌های متناظر با داده‌های دورافتاده در ماتریس‌های X_{L-1} و Y_{L-1} را محاسبه نماییم. تجزیه مقدار تکین رابطه $X_{L-1} = U_X D_X V_X' = U_X^* D_X^* V'^*_X$ انجام شده و ماتریس‌های $X_{L-1}^* = X_{L-1}V_X^* = U_X^* D_X^*$ و $Y_{L-1}^* = Y_{L-1}V_Y^* = U_Y^* D_Y^*$ در نظر گرفته می‌شوند. وزن‌های $w_i(X_{L-1}^*)$ برای سطر دورافتاده $x_i^{*(L-1)}$ ($i = 1,\ldots,n$) از ماتریس X_{L-1}^* تعریف می‌شود(کراکس و همکاران، ۲۰۰۳):

$$w_i(X_{L-1}^*) = \min\left(1, \frac{\chi^2_{p^*,0/95}}{RD_i^2(X_{L-1}^*)}\right) \quad , \quad i = 1,\ldots,n$$

در این‌جا، $\chi^2_{p^*,0/95}$ مقدار بحرانی ۵% بالایی توزیع کای‌دو با p^* درجه آزادی می‌باشد که $p^* = k - L + 1$ تعداد ستون‌های X_{L-1}^* می‌باشد و برای $i = 1,\ldots,n$

$$RD_i\left(X_{L-1}^*\right) = \sqrt{\left(x_i^{*(L-1)} - T\left(X_{L-1}^*\right)\right)' C\left(X_{L-1}^*\right)^{-1}\left(x_i^{*(L-1)} - T\left(X_{L-1}^*\right)\right)}$$

فاصله استوار می‌باشد (روسیو و ون زومرن[۱۰]، ۱۹۹۰). برآوردگرهای چندمتغیره استوار پراکنشی و مقایسی T و C، به عنوان قسمت مقایسی و پراکنشی برآوردگر MVE (روسیو، ۱۹۸۵) محاسبه شده از X_{L-1}^* انتخاب می‌شوند. برآوردگر MVE به این دلیل

[۸] Robust Alternating Regressions
[۹] Least Trimmed Squares
[۱۰] Rousseeuw & Van Zomeren

این جا انتخاب می‌شود که به عنوان یک شناسه دورافتاده، به خوبی ایفای نقش می‌کند (بکر و گتر[۱۱] ، ۲۰۰۱).

۳ شبیه‌سازی

با استفاده از نرم‌افزار S-plus و معیار میانگین توان دوم خطا (MSE) و با $n = ۵۰۰$ مشاهده و تعداد $nsimul = ۱۰۰۰$ شبیه‌سازی را انجام می‌دهیم. توزیع‌های نمونه زیر را در نظر می‌گیریم:

-- توزیع نرمال (NOR) با میانگین صفر و ماتریس کوواریانس Σ، $N_{p+q}(۰, \Sigma)$.
-- توزیع t چندمتغیره با ۳ درجه آزادی.
-- آلودگی‌متقارن (SCN): توزیع $N_{p+q}(۰, \Sigma)$ با آلودگی ۵% از توزیع $N_{p+q}(۰, ۹\Sigma)$.
-- آلودگی نامتقارن (ACN): $N_{p+q}(۰, \Sigma)$ با آلودگی ۵% از داده‌های معادل نقطه $tr(\Sigma)۱'$.

برای پارامترهای برآورد شده $(L = ۱, \ldots, k)$ $\hat{a}_L^j$، $\hat{b}_L^j$ و $\hat{\rho}_L^j$ داریم

$$MSE(\hat{a}_L) = \frac{۱}{nsimul} \sum_{j=۱}^{nsimul} \cos^{-۱}\left(\frac{|a'_L \hat{a}_L^j|}{||\hat{a}_L^j|| \cdot ||a_L||}\right),$$

$$MSE(\hat{b}_L) = \frac{۱}{nsimul} \sum_{j=۱}^{nsimul} \cos^{-۱}\left(\frac{|b'_L \hat{b}_L^j|}{||\hat{b}_L^j|| \cdot ||b_L||}\right),$$

$$MSE(\hat{\rho}_L) = \frac{۱}{nsimul} \sum_{j=۱}^{nsimul} \left(\phi(\hat{\rho}_L^j) - \phi(\rho_L)\right)^۲$$

که $\phi(\rho_L) = \tanh^{-۱}(\rho_L)$ تبدیل فیشر ρ_L می‌باشد (برانکو و همکاران، ۲۰۰۵).

در همه شکل‌ها (به علت این که تعداد شکل‌ها بسیار زیاد بود، از آوردن شکل‌ها خودداری شده است)، واضح است که شبیه‌سازی مربوط به آلودگی نامتقارن (ACN) به بزرگ‌ترین MSE منجر شده و برآوردگر روش کلاسیک بدتر از روش‌های دیگر می‌باشد. برای متغیرهای متعارف مرتبه اول، بهترین برآوردگر، برآوردگر ضریب همبستگی رتبه‌ای اسپیرمن و برای متغیرهای متعارف مراتب بالاتر، بهترین برآوردگر، برآوردگر MCD می‌باشد. این درحالی است که روش رگرسیون متناوب استوار نتایج خوبی را نشان نمی‌دهد. به‌طورمیانگین (هندسی)، همبستگی متعارف مرتبه اول درصدی

[۱۱] Becker & Gather

از همبستگی کل بین X و Y را شامل می‌شود که درصد مذکور برای روش‌های مختلف، در بازه ٦۴% تا ۷۷% قرار دارد.

برای بُعدهای $p = ۲$ و $q = ۴$ و همچنین بُعدهای $p = ۴$ و $q = ۴$ نتایج مشابهی به دست می‌آید. آلودگی نامتقارن (ACN) به بزرگ‌ترین MSE در مقایسه با $T۳$ و SCN منجر می‌شود. همانند قبل، روش کلاسیک به بدترین وضعیت انجام می‌شود. افزایش RAR در بردارهای متعارف مراتب بالاتر، مجدداً قابل مشاهده می‌باشد. نتایج خیلی خوبی با روش تعقیب تصویر بر اساس ضریب همبستگی رتبه‌ای اسپیرمن به دست می‌آید. به طور میانگین (هندسی)، درصد همبستگی متعارف مرتبه اول از همبستگی کل در بازه ۵۰% تا ٦٦% ، برای مرتبه اول و دوم در بازه ۷۷% تا ۸٦% و برای مرتبه اول، دوم و سوم در بازه ۸۹% تا ۹۷% قرار دارند.

اگر برای بُعدهای $p = ۲$ و $q = ۲$ مجدداً شبیه‌سازی تحت این شرایط که ε درصد (از صفر تا ۲۵ درصد) از داده‌ها آلوده باشند، انجام شود، ملاحظه می‌شود که در روش کلاسیک، MSE در حضور آلودگی، به سرعت افزایش می‌یابد. روش تعقیب تصویر بر اساس برآوردگر MCD پایدار می‌باشد که تا حدود ۲۰% آلودگی، MSE کوچک باقی می‌ماند ولی بعد از آلودگی ۲۰% افزایش می‌یابد. MSE برای روش رگرسیون متناوب استوار (RAR) با افزایش آلودگی، به طور یکنواخت در حال افزایش است. مطلوب‌ترین حالت، روش تعقیب تصویر بر اساس همبستگی رتبه‌ای اسپیرمن ($PP-SPM$) می‌باشد. $PP-M$ یک رفتار نسبتاً غیرمترقبه نشان می‌دهد؛ برای اولین همبستگی متعارف در کل دامنه آلودگی، MSE خیلی کوچک باقی می‌ماند ولی در همبستگی‌های متعارف مراتب بالاتر، از حدود آلودگی ۱۵% به بعد روند افزایشی را مشاهده می‌کنیم.

۴ بحث و نتیجه‌گیری

روش تعقیب تصویر بر اساس همبستگی رتبه‌ای اسپیرمن به وضوح به MSE کوچک‌تری منجر شده و ارجح‌ترین برآوردگر تعقیب تصویر مطرح شده می‌باشد. در همه آزمایش‌های شبیه‌سازی، مشاهده شد که الگوریتم RAR استوار است ولی در کل، نتایج نسبتاً ضعیف‌تری نسبت به دیگر روش‌ها ارائه نمود.

مراجع

Becker, C. and Gather, U. (2001), *The largest nonidentifiable outlier: A comparison of multivariate simultaneous outlier identification rules*, Computational Statistics and Data Analysis.

Branco J.A., Croux C., Filzmoser P. and Oliveira M.R. (2005), *Robust Canonical Correlations: A Comparative Study*, Computational Statistics.

Croux, C. and Dehon, C. (2000), *Robust Canonical Correlations using High Breakdown Scatter Matrices*, Preprint, University Libre de Bruxelles.

Croux, C. and Dehon, C. (2002), *Analyse canonique basee sur des estimateurs robustes de la matrice de covariance*, La Revue de Statistique Appliquge.

Croux, C., Filzmoser, P., Pison, G. and Rousseeuw, P.J. (2003), *Fitting multiplicative models by robust alternating regressions*, Statistics and Computing.

Huber, P.J. (1981), *Robust Statistics*, John Wiley.

Johnson, R. A. and Wichern, D. W. (2000), *Applied multivariate statistical analysis*, Prentice Hall, New Jersey, 4th edition.

Karnel, G. (1991), *Robust canonical correlation and correspondence analysis*, In: the Frontiers of statistical Scientific and Industrial Applications, (Volume II of the proceedings of ICOSCO-I, The First International Conference on Statistical Computing), American Sciences Press, Strassbourg.

Maronna, R.A. (1976), *Robust M-estimators of multivariate location and scatter*, The Annals of Statistics.

Rousseeuw, P.J. (1984), *Least median of squares regression*, Journal of the American Statistical Association.

Rousseeuw, P.J. (1985), *Multivariate estimation with high breakdown point*, In: W. Grossmann, G. Pflug, I. Vincze, and W. Wertz (eds.), Mathematical Statistics and Applications, Vol. B, Reidel, Dordrecht.

Rousseeuw, P.J. and Van Driessen, K. (1999), *A fast algorithm for the minimum covariance determinant estimator*, Technometrics.

Rousseeuw, P.J. and Van Zomeren, B. (1990), *Unmasking multivariate outliers and leverage points*, Journal of the American Statistical Association.

Wold, H. (1966), *Nonlinear estimation by iterative least squares procedures*, In: F.N. David (ed.), A Festschrift for J. Neyman, Wiley, New York.