تحليل المركبات الرئيسة الحصينة بالاعتماد على بعض طرائق متابعة الاسقاطات

أ.م.د. قتيبة نبيل نايف القزاز/ جامعة بغداد/ كلية الادارة والاقتصاد
م.د. رباب عبد الرضا صالح/ جامعة بغداد/ كلية الادارة والاقتصاد

المستخلص

أن تحليل المركبات الرئيسة التقليدية تكون حساسة تجاه القيم الشاذة اذ يتم حسابها من خلال القيم والمتجهات المميزة لمصفوفة التباين أو الأرتباط غير الحصين والتي تعطي نتائج غير صحيحة في حالة أحتواء البيانات على القيم الشاذة. ولمعالجة هذة المشكلة يتم اللجوء الى أستعمال الطرائق الحصينة، اذ توجد عدة طرائق حصينة سوف يتم التطرق الى بعض منها. أن مقدرات القياس الحصينة تتضمن حساب المقدرات الحصينة مباشرة للقيم المميزة بدون الأعتماد على المقدرات الحصينة لمصفوفة التباين والتباين المشترك. كذلك أن تحليل المركبات الرئيسة يبحث عن سلوك أعلى أنتشار للبيانات المسقطة على الأتجاهات لكن بدلا من أستعمال التباين كمقياس للأنتشار سيتم أستعمال مقدرات القياس الحصينة كمؤشر لمتابعة الأسقاطات Projection-Pursuit.

في هذا البحث تم أستعمال خوارزمية Croux and Ruiz-Gazen حيث أن المركبات الرئيسة الحصينة تعرف بأسقاطات البيانات على أعلى أتجاه لمقدر القياس الحصين، اذ سيتم التركيزعلى القياس الحصين الى (MAD و Qn).

المصطلحات الرئيسية للبحث/ تحليل المركبات الرئيسة – متابعة الاسقاطات – الحصانة.

مجلة العلوم الاقتصادية والإدارية
المجلد ٢١ العدد ٨٣
الصفحات ٣١٧-٣٢٧

١-١المقدمة

تحليل المركبات الرئيسة PCA غالبا مايكون الخطوة الاولى في تحليل البيانات اذ يليه التحليل باستعمال انحدار المركبات الرئيسة PCR والمربعات الصغرى الجزئية PLS والتحليل المميز وغيرها من تقنيات متعدد المتغيرات [8] .

ويتضمن تحليل المركبات الرئيسة PCA حساب المتجهات و القيم المميزة لمصفوفتي التباين او الارتباط الا ان تحليل المركبات الرئيسة غالبا ما يكون حساس تجاه القيم الشاذة وهناك طرائق حصينة تستعمل لغرض حصانة المركبات الرئيسة والتي يمكن تصنيفها على النحو الاتي :

الطريقة الاولى تعتمد على استبدال مصفوفة التباين التقليدية الى مصفوفة التباين الحصينة بالاعتماد على مقدري الموقع والتباين الحصينين ومن ثم حساب المتجهات المميزة والقيم المميزة بالاعتماد على التقديرات الحصينة لمصفوفة التباين والتباين المشترك اذ درس الباحث (Devlin et al. 1981) مقدر M لمصفوفة التباين والتباين المشترك اذ اعطى للمقدر الحصين نقطة انهيار منخفضة الابعاد وان نقطة الانهيار العالية للمقدر الحصين لمصفوفة التباين والتباين المشترك تم الحصول عليه من الباحث Rousseeuw (1985) للمقدر (Minimum Volume Ellipsoid) وغيرها من الطرائق اما الطريقة الثانية تدعى بطريقة متابعة الاسقاطات وتوجد عدة خوارزميات منها تعتمد على فكرة متابعة الاسقاطات التي تتضمن حساب المقدرات الحصينة للمتجهات والقيم المميزه بشكل مباشر وبدون عملية تحصين مقدر مصفوفة التباين والتباين المشترك وان طريقة متابعة الاسقاطات (pp(Projection- Pursuit طورت من قبل الباحثان (Li and Chen 1985) اذ كان الباحث (Huber)[3,5]قد اقترحها عام ١٩٨٥ ايضا.

اذ ان تحليل المركبات الرئيسية التقليدية PCA يبحث على الاتجاهات لاعلى انتشار للبيانات المسقطة على الاتجاهات لكن بدلا من استعمال التباين كمقياس للانتشار سيستعمل مقدرات القياس الحصينة كمثال استعمال MAD (median absolute deviation) و Sn (QN, scale estimator of the first quartile كمؤشر لمتابعة الاسقاطات Projection- Pursuit اما الطريقة الثالثة فتجمع بين الطريقة الاولى والثانية وهذه الطريقة تعالج الابعاد العالية للبيانات مثل طريقة ROBPCA. اما الطريقة الرابعة فتعد كطريقة معدلة تعمل على تعديل الحسابات الداخلة في خوارزمية تجزئة القيم المفردة Singular Value Decomposition(SVD) عن طريق استبدال معيار المربعات الصغرى مع مقدر حصين وهذا يمكن التعامل معه للابعاد العليا والقيم المتطرفه [8]

١-٢ هدف البحث

ان الهدف من البحث هوالمقارنة بين مقدر القياس الحصين MAD Median Absolute Deviation ومقدر القياس الحصين Qn, scale estimator of the first quartile من خلال قياس دقة المقدر بالنسبة الى الانحراف القياسي standard deviation STD للقيم المميزة λ ومؤشر الاسقاطات لمقدرات القياس الحصينة Qn, MAD للقيم المميزة لتجارب محاكاة لعدة انواع من القيم ولابعاد مختلفة لبيانات تتضمن قيم شاذة وبيانات لاتتضمن قيم شاذة وعندما يتوزع حد الخطأ توزيعا طبيعياً اووفق توزيع لابلاس .

١-٣ مشكلة البحث

ان من المشاكل التي تواجه الباحث عند تحليل البيانات هو وجود التلوث في البيانات ولاسيما بيانات متعدد الابعاد اذ تؤثر هذه المشكلة على مصفوفة التباين والتباين المشترك ومنه مصفوفة الارتباط مما يؤدي عند استعمال تحليل المركبات الرئيسة الكلاسيكية الى عدم دقة النتائج في تفسير العلاقة بين المتغيرات قيد الدراسة وعلية تمت معالجة وجود التلوث باستعمال المركبات الرئيسة الحصينة بالاعتماد على مقدرات القياس الحصينة كمؤشر لمتابعة الاسقاطات والمتمثلة بمقدري القياس الحصينين MAD,Qn ومن ثم المقارنة بين كفاءة هذين المقدرين من خلال تجارب المحاكاة .

٢- الجانب النظري

٢-١ تحليل المركبات الرئيسة

ان طريقة المركبات الرئيسة تهدف الى تقليص البيانات وتصف التباين للبيانات عن طريق انشاء مجموعة جديدة من المركبات المتعامدة تسمى بالمركبات الرئيسة وهذه المركبات هي تراكيب خطية من المتغيرات وتكون متعامدة بعضها مع البعض وكل مركبة تصف جزءا من التباين للبيانات ولاتفسر التي تسبقها ولهذا قد تكون المركبة الرئيسة الاولى تكفي لتمثيل معظم تباين البيانات حيث يتم تحليل البيانات الاصلية بواسطة تحليل المركبات الرئيسة PCA الى مصفوفة القياس scores matrix ويرمز لها بـ T ومصفوفة التحميل loading matrix ويرمزلها بـ P حيث ان X= TP' حيث ان X مصفوفة البيانات data matrix.

ان اهمية المركبات الرئيسة لا تتمثل بتقليص البيانات فقط وانما يتم استعمالها كخطوة اولى في تحليل البيانات لمتعدد المتغيرات فيتم استعمالها في الانحدار والتحليل المميز.

وهذا يؤدي الى معالجة مشكلة المتغيرات المترابطة والمصفوفة المفردة او القريبة من المفردة لـ X'X ويتم التغلب على هذه المشكلة بالاستعاضة عن المتغيرات بالمركبات الرئيسية المتعامدة وهي نفس البيانات الاصلية الاان هذه المركبات تكون عرضة للقيم المتطرفه ولهذا نلجأ الى استعمال الطرائق الحصينة لمعالجة القيم الشاذة في البيانات [4].

٢-٢ تحليل المركبات الرئيسة الحصينة

ان نتائج تحليل المركبات الرئيسة التقليدية PCA تتاثر بالقيم المتطرفة وهناك عدة طرائق حصينة للمركبات الرئيسة لمعالجة وجود القيم الشاذة واحدى هذه الطرائق التي تم الاعتماد عليها في هذا البحث هي طريقة متابعة الاسقاطات ومن الخوارزميات التي تستعمل في هذه الطريقة هي خوارزمية Croux and Ruiz- Gazen.

ففي تحليل المركبات الرئيسة الحصينةRPCA نبحث عن الاتجاه لانتشار البيانات المسقطة عوضا عن استعمال التباين التقليدي يتم استعمال المقياس الحصين Sn كمؤشر لمتابعة الاسقاط لسلسلة من المشاهدات المتتابعة $X1...Xn \in IR^p$ اول متجه مميز يعرف كالاتي [2,3]

$$V_{Sn,1} = \arg\max_{\|a\|=1} Sn(a'x1,\ldots,a'xn) \quad .... \qquad (2-1)$$

والقيمة المميزه تعرف كالاتي

$$\lambda_{Sn,1} = S_n^2((V_{Sn,1})'x1,\ldots,(V_{Sn,1'})'xn) \quad ... \qquad (٢-٢)$$

ولـ kthمن المتجهات الميزة والقيم المميزة تعرف كالاتي :

$$V_{Sn,k} = \arg\max_{\|a\|=1, a \perp Vsn,1,\ldots, a \perp Vsn,(k-1)} Sn(a'x1,\ldots,a'xn) \quad ... \quad (2-3)$$

$$\lambda_{Sn,n} = S_n^2((V_{Sn,k})'x1,\ldots,(V_{Sn,k})'xn) \quad \ldots(2-4)$$

اما المركبات الرئيسة الحصينة فتعطى بواسطة اسقاطات المشاهدات على المتجه المميز وفي ادناه وصف لخوارزمية Croux and Ruiz- Gazen.

٢-٣ وصف لخوارزمية Croux and Ruiz- Gazen

١- نعمل تمركز لمصفوفة البيانات $X_{(m,n)}$ حول الوسيط المكاني (وهومقدر حصين للموقع في متعدد المتغيرات وله خصائص حصينة جيدة تتمثل بخاصية التغاير المتساوي ويكون المقدر متعامد) للحصول على X_C ونحدد الرتبة r وكما يأتي [4,6,7] :

$$X_C = X - I'1_{l1}(X) \quad ... \qquad (2-1)$$

حيث ان : r = min (m−1, n)

٢- بناء P_i من المتجهات تكون كصفوف طبيعية من المصفوفة X_{ci} وكما يأتي :

$$P_i = \frac{X_{ci}}{\|X_{ci}\|}; \quad i=1,2,\ldots,m \qquad ... \qquad (2-2)$$

٣- اسقاط جميع المشاهدات على الاتجاهات الممكنة وكما يأتي :

$$t_i = X_c \; p'_i \qquad ... \qquad (2-3)$$

٤- نختار احد المقدرات الحصينة للتباين $\sigma_S(t_i)$ حيث S= MAD or Qn or Sn

$$\sigma_{MAD}(t_i) = C.median_i \left| t_i - median_j(t_j) \right| \quad ...(٢-٤)$$

حيث C ثابت ويساوي 1.4826

$$\sigma_{Sn}(t_i) = cf.C_{sn}.median_i\{median_j \left| t_i - t_j \right| \quad ... \quad (2-5)$$

حيث cf هو عامل الاتساق يعتمد على حجم البيانات و C_{sn} ثابت ويساوي 1.1926

$$\sigma_{Qn}(t_i) = cf.C_{qn}.\{\left| t_i - t_j ; \quad 1 < j \right|\}_{(K)} \quad ... \qquad (٢-٦)$$

حيث cf هو عامل الاتساق يعتمد على حجم البيانات و C_{qn} ثابت ويساوي 2.2219

$$K = \binom{h}{2} = \binom{n}{2}/4 \quad where \quad h = [n/2]+1$$ و

٥- ايجاد الاتجاه v الذي يعظم مصفوفة $\sigma_S(t_i)$ للاسقاطات وكما يأتي :

$$V = \max(\sigma_S(t_i)) \qquad ... \qquad (2-7)$$

٦- اسقاط كل المشاهدات على الاتجاه المختار V للحصول على المركبة الرئيسه الحصينة وكما يأتي :

$$t_i = X_C P'_V \quad ... \qquad (٢-٨)$$

٧- تجديد مصفوفة البيانات بواسطة المكمل المتعامد وكما يأتي :

$$X_c = X_c - (PvP'v)X_c \quad ... \qquad (٢-٩)$$

٨- اذا كان عدد المركبات الرئيسية الحصينة اقل من r نرجع للخطوة ٢

٣-١ الجانب التجريبي

من اجل الوصول الى الهدف من البحث تم استعمال اسلوب المحاكاة لمقارنة بين طريقتي المكونات الرئيسية والمكونات الرئيسة الحصينة وباستعمال الكفاءة الى خوارزمية Croux–Ruiz (CR) نسبة الى خوارزمية PCA لابعاد متغيرات وحجوم عينات مختلفة ولنسب تلوث مختلفة وقد تم كتابة البرنامج بلغة ماتلاب (version R2011).

٣-٢ وصف تجربة المحاكاة

تم توليد البيانات بحجوم عينات (n= ٢٥,٥٠,٧٥,١٠٠) وعدد المتغيرات التوضيحية (P=5,10,20)) وقد تم تكرار التجربة ١٠٠٠ مرة وتمت المقارنة بحسب المقياس الاتي [2]

$$EFF_1(CR(MAD),PCA)=\frac{\hat{\lambda_1}_{CR(MAD)}}{\hat{\lambda_1}_{PCA}} \qquad \cdots(3-1)$$

$$EFF_1(CR(Qn),PCA)=\frac{\hat{\lambda_1}_{CR(Qn)}}{\hat{\lambda_1}_{PCA}} \qquad \cdots(3-2)$$

حيث ان $\hat{\lambda}_{1CR(MAD)}$ و $\hat{\lambda}_{1CR(Qn)}$ تشير الى القيم المميزة الاولى بطريقة Croux–Ruiz (CR) لمقدري القياس الحصينين (MAD و Qn).

$\hat{\lambda}_{1\,PCR}$ تشير الى القيمة المميزة الاولى بطريقة المركبات الرئيسة .

وبالاعتماد على نموذج معادلة الانحدارالخطي المتعدد وكما يأتي :

$$y=X\beta+\varepsilon \qquad \ldots \qquad (3-3)$$

حيث ان :-

y : موجة المتغير المعتمد ببعد nx1

X : مصفوفة المتغيرات التوضيحية ذات رتبة nxp

β : موجه معلمات معادلة الانحدار غير المعلومة ببعد px1

ε : موجه الاخطاء العشوائية ببعد nx1

تم توليد البيانات بالصيغة الاتية [1] :

$$\begin{aligned} x_1 &= N(0,1) \\ x_{p-1} &= N(0,0.1)+x_1 \\ y &= x_1+\cdots+x_p+N(0,1) \end{aligned} \qquad \cdots(3-4)$$

وتم توليد حد الخطأ بالتوزيع الطبيعي القياسي N(0,1) وتوزيع لابلاس

كما تم توليد البيانات الملوثة كما يأتي:

$$0.9\ N(0,1) + 0.1\ N(15,1)$$

$$0.8\ N(0,1) + 0.2\ N(15,1)$$

$$0.7\ N(0,1) + 0.3\ N(15,1)$$

$$0.6\ N(0,1) + 0.4\ N(15,1)$$

٤ - نتائج المحاكاة

تم تقدير القيم المميزة للطرائق المذكورة في الجانب النظري ولجميع العينات ومختلف اعداد المتغيرات التوضيحية والتي تم ذكرها كما في الجداول الاتية :

جدول رقم (١) معدل الكفاءة الى خوارزمية MAD, Q نسبة الى خوارزمية PCA للبيانات الخالية من التلوث

P / N	5		١٠		٢٠	
	MAD	Q	MAD	Q	MAD	Q
25	0.2293	0.2946	0.4560	0.6515	0.5547	0.8759
٥٠	0.2259	0.2686	0.4261	0.4628	0.4784	0.5946
٧٥	0.2017	0.2425	0.4079	0.4266	0.4302	0.5874
١٠٠	0.1724	0.2299	0.2851	0.3315	0.2983	0.4319

جدول رقم (٢) معدل الكفاءة الى خوارزمية MAD, Q نسبة الى خوارزمية PCA للبيانات الملوثة بنسبة ١٠%

P / N	5		١٠		٢٠	
	MAD	Q	MAD	Q	MAD	Q
25	0.7549	0.7934	0.5650	0.7838	0.5810	0.7371
٥٠	0.3043	0.3083	0.4673	0.5034	0.5605	0.6436
٧٥	0.2481	0.2574	0.2850	0.3289	0.4916	0.5391
١٠٠	0.2020	0.2239	0.2697	0.2797	0.2801	0.3048

جدول رقم (٣) معدل الكفاءة الى خوارزمية MAD, Q نسبة الى خوارزمية PCA للبيانات الملوثة بنسبة ٢٠%

P / N	5		١٠		٢٠	
	MAD	Q	MAD	Q	MAD	Q
25	0.3515	0.6071	0.5766	0.5970	0.7237	0.7362
٥٠	0.3430	0.4503	0.3928	0.5794	0.5270	0.7523
٧٥	0.2127	0.2565	0.3035	0.3832	0.3735	0.4759
١٠٠	0.2١٣٩	0.2٣٠٢	0.٣٣٧٤	0.٣٥٦٩	0.3079	0.4298

جدول رقم (٤) معدل الكفاءة الى خوارزمية MAD, Q نسبة الى خوارزمية PCA للبيانات الملوثة بنسبة ٣٠%

P / N	5		١٠		٢٠	
	MAD	Q	MAD	Q	MAD	Q
25	0.2967	0.3449	0.3329	0.5010	0.4629	0.7249
٥٠	0.2892	0.3085	0.4580	0.4754	0.4901	0.5611
٧٥	0.2253	0.2325	0.2730	0.3410	0.4413	0.4976
١٠٠	0.2070	0.2410	0.2228	0.2485	0.3637	0.3890

جدول رقم (٥) معدل الكفاءة الى خوارزمية MAD, Q نسبة الى خوارزمية PCA للبيانات الملوثة بنسبة ٤٠ %

P / N	5		١٠		٢٠	
	MAD	Q	MAD	Q	MAD	Q
25	0.4408	0.6114	0.6704	0.7441	0.8686	0.9725
٥٠	0.2270	0.2994	0.3738	0.3928	0.8038	0.8606
٧٥	0.2465	0.2626	0.3246	0.3984	0.4490	0.5963
١٠٠	0.2205	0.2242	0.2751	0.3492	0.4232	0.4412

جدول رقم (٦) معدل الكفاءة الى خوارزمية MAD, Q نسبة الى خوارزمية PCA للبيانات التي يتوزع الخطا laplace

P / N	5		١٠		٢٠	
	MAD	Q	MAD	Q	MAD	Q
25	0.4564	0.5802	0.5506	0.5796	0.5284	0.8247
٥٠	0.2861	0.3432	0.4074	0.4848	0.4622	0.5991
٧٥	0.2788	0.2884	0.2907	0.3693	0.3283	0.5104
١٠٠	0.2141	0.2273	0.2710	0.3452	0.3341	0.4016

٥- الاستنتاجات

من نتائج تجارب المحاكاة تم التوصيل الى الاستنتاجات الاتية:

١- في حالة حد الخطا يتوزع توزيع طبيعي وباستعمال المعيارين $EFF_1(CR(MAD),PCA)$ و $EFF_1(CR(Qn),PCA)$ وجدنا ان الطريقة الحصينة CR(MAD) هي اكثر كفاءة من الطريقة التقليدية PCA وان الطريقة الحصينة CR(Qn) هي اكثر كفاءة من طريقة التقليدية PCA وان الطريقة الحصينة CR(MAD) اكثر كفاءة من الطريقة الحصينة CR(Qn) لجميع الحجوم والابعاد ونجد ان الكفاءة تقل بازدياد عدد المتغيرات وتزاد الكفاءة بزيادة المشاهدات.

٢- في حالة تلويث حد الخطا بنسبة ١٠ % و٢٠ % و ٣٠ % و٤٠ % وباستعمال المعيارين $EFF_1(CR(MAD),PCA)$ و $EFF_1(CR(Qn),PCA)$ وجدنا ان الطريقة الحصينة CR(MAD) هي اكثر كفاءة من الطريقة التقليدية PCA وان الطريقة الحصينة CR(Qn) هي اكثركفاءة من طريقة التقليدية PCA وان الطريقة الحصينة CR(MAD) هي اكثر كفاءة من الطريقة الحصينة CR(Qn) لجميع الحجوم والابعاد ونجد ان الكفاءة تقل بازدياد عدد المتغيرات وتزاد الكفاءة بزيادة المشاهدات .

٣- في حالة حد الخطا يتوزع لابلاس وباستعمال المعيارين $EFF_1(CR(MAD),PCA)$ و $EFF_1(CR(Qn),PCA)$ وجدنا ان الطريقة الحصينة CR(MAD) هي اكثر كفاءة من الطريقة التقليدية PCA وان الطريقة الحصينة CR(Qn) هي اكثر كفاءة من طريقة التقليدية PCA وان الطريقة الحصينة CR(MAD) هي اكثر كفاءة من الطريقة الحصينة CR (Qn) لجميع الحجوم والابعاد ونجد ان الكفاءة تقل بازدياد عدد المتغيرات وتزاد الكفاءة بزيادة المشاهدات.

٦- التوصيات

بناء" على ماتم التوصل اليه من استنتاجات في الجانب التجريبي يمكن ادراج التوصيات الاتية التي يراها الباحث ضروريه ، لقد كانت الطريقتان الحصينتان MAD و Qn اكثر دقة من الطريقة التقليدية PCA وكانت الطريقة الحصينة MAD اكثر دقة من الطريقة الحصينة Qn ولجميع حجوم العينات والابعاد المختلفه ونسب شواذ مختلفة لذا نوصي باستعمال الطريقة الحصينة MAD في حالة حد الخطا يتوزع توزيعا طبيعيا وغير طبيعي (توزيع لابلاس) وفي حالة البيانات الملوثة ولنسب تلوث مختلفة.

المصادر

1- Adnan,N., (2006) "A comparative Study on Methods For Handling Multicollinearity Problems" J. Matematika, V. 22, N.2, PP. 109_119

2-Croux,C., Filzmoser,P. , Oliveira,M.R. (2007) "Algorithms for Projection-Pursuit Robust Principal Component Analysis"J.Chemometrics and Intelligent Laboratory Systems ,V87,Issue 2. PP 218-225

3- Croux,C.,Ruiz-Gazen,A.,(2005) ''High Breakdown Estimators for Principal Components: the Projection-Pursuit Approach Revisited'' Journal of Multivariate Analysis, V 95,Issue 1,pP 206-226

4-Daszykowski,M.,Kaezmarek,K.,Heyden,Y.,Walczak,B .(2007) "Robust statistics in data analysis —A review Basic concepts" Chemometrics and Intelligent Laboratory Systems,V 85 ,Issue 2, PP 203–219

5- Filzmoser ,P.,Fritz,H.(2007)"Exploring high-dimensional data with robust principal components "Technische Universitat Wien,PP 1-8

6-Gharibnezhad,F., Mujica,L. ,Rodellar,.J(2011)"Comparison of two robust PCA methods for damage detection in presence of outliers " Journal of Physics: Conference Series Voume 305 (2011) 012009,doi:10,1088/1742-6596/305/1/01/2009, IOP Publishing

7- Hubert, M., Rousseeuw, P.J., Verboven,S (2001), ''A fast robust method for principal components with applications to chemometrics" J . Chemometrics and Intelligent Laboratory Systems, 60,101-111

8- Moller,S., Frese,J., Bro,R.,(2005)''Robust methods for multivariate data analysis" Journal of Chemometrics, Volume 19, Issue 10, pages 549–563, October 2005

Analysis of Robust Principal Components Depends on the some methods of Projection-Pursuit

Abstract

The analysis of the classic principal components are sensitive to the outliers where they are calculated from the characteristic values and characteristic vectors of correlation matrix or variance Non-Robust, which yields an incorrect results in the case of these data contains the outliers values. In order to treat this problem, we resort to use the robust methods where there are many robust methods Will be touched to some of them.

The robust measurement estimators include the measurement of direct robust estimators for characteristic values by using characteristic vectors without relying on robust estimators for the variance and covariance matrices. Also the analysis of the principal components search for the trends of the highest scattered data projected on these vectors, but instead of using the variance as a measure for scattering, we will use robust measurement estimators as indicator for Projection-Pursuit.

In this paper, we used Croux and Ruiz-Gazen algorithm, where the principal components are recognize by projection data on the highest vector for robust measurement estimators, focusing on the robust measurement to Qn and MAD.

Keywords: **Robust Principal Components, Projection-Pursuit and Robustness.**